

Université de Montréal

Impact épigénomique de mutations associées à des syndromes neurodéveloppementaux dans
des régulateurs de la chromatine

par Sophie Ehresmann

Programmes de Biologie Moléculaire
Faculté de Médecine

Mémoire présenté en vue de l'obtention du grade de Maîtrise
en Biologie Moléculaire
option Générale

Avril, 2019

© Sophie Ehresmann, 2019

Résumé

Les syndromes neurodéveloppementaux, caractérisés par un déficit intellectuel, un retard global du développement, ou un trouble du spectre de l'autisme, affectent environ cinq pourcents de la population. De ceux-ci, une grande partie de leur génétique sous-jacente est encore inconnue. Des mutations dans des composantes de complexes de remodelage de la chromatine ont été associées précédemment à plusieurs syndromes neurodéveloppementaux, impliquant des modifications épigénétiques de l'ADN ou d'histones, et le remodelage de nucléosomes. Nous avons identifié des individus porteurs de mutations dans *TRRAP*, un gène impliqué dans les complexes histone acétyl-transférase, *SMARCC2*, une protéine structurale du complexe de remodelage des nucléosomes SWI/SNF, et *CHD3*, l'une des hélicases ATP-dépendantes du complexe de remodelage des nucléosomes NuRD. Pour les trois, nous avons étudié à l'échelle cellulaire l'expression génique globale et l'ouverture de la chromatine dans des cellules provenant de certains de ces individus par séquençage à haut débit. Ainsi pour chacun des groupes d'individus, nous avons pu démontrer que certains gènes sont différenciellement exprimés, et dans le cas d'individus porteurs de mutations dans *TRRAP*, que plusieurs de ces gènes sont associés à une plus grande ouverture de la chromatine au niveau de leur promoteur; une grande partie d'entre eux ayant un rôle dans le développement neural ou la fonction neuronale.

Mots-clés

Syndrome Neurodéveloppemental, RNAseq, ATACseq, Acétylation d'histones, remodelage de nucléosomes, *TRRAP*, *SMARCC2*, *CHD3*

Abstract

Neurodevelopmental syndromes, characterized by a global developmental delay, intellectual deficiency, or autism spectrum disorders, affect around five percent of the population. Of these, a large part of their genetic causes remains unresolved. Lately, mutations in subunits of large chromatin remodeling complexes have been associated with neurodevelopmental syndromes, affecting DNA and histone modifications, as well as nucleosome remodeling. We have identified individuals carrying mutations in *TRRAP*, a component of multiple histone acetyltransferase complexes; *SMARCC2*, a structural protein involved in the formation of the BAF (SWI/SNF) nucleosome remodeling complex; and *CHD3*, one of the ATP-dependent helicases of the NuRD nucleosome remodeling complex. For these three genes, we characterized gene expression patterns and chromatin structure for several individuals harboring *de novo* mutations by RNAseq and ATACseq in individuals' fibroblasts or lymphoblastoid cell lines. We observed that they showed differential expression for a number of genes, of which a large proportion is associated with neural development or neuronal function. We also showed that some differentially expressed genes in *TRRAP* individuals are also associated with a higher chromatin opening at their promoter.

Keywords

Neurodevelopmental syndrome, RNAseq, ATACseq, Histone acetylation, Nucleosome remodeling, *TRRAP*, *SMARCC2*, *CHD3*

Table des matières

| | |
|--|------|
| Université de Montréal | ii |
| Résumé..... | i |
| Mots-clés..... | i |
| Abstract | ii |
| Keywords | ii |
| Table des matières..... | iii |
| Liste des tableaux..... | vi |
| Liste des figures | vi |
| Liste des sigles | viii |
| Remerciements..... | xi |
| 1. Introduction..... | xi |
| 1.1 Remarques générales | 1 |
| 1.2 Les mécanismes de l'épigénétique..... | 2 |
| 1.2.1 L'ADN et sa composition | 2 |
| 1.2.2 La chromatine et les nucléosomes | 4 |
| 1.2.3 Les histones et leurs variantes..... | 5 |
| 1.2.4 Les modifications d'histones | 7 |
| 1.3. Le développement neural | 9 |
| 1.3.1 Différenciation séquentielle des différents types cellulaires du cerveau | 9 |
| 1.3.1.a Cellules Souches Neurales et Cellules Neuroépithéliales | 9 |
| 1.3.1.b Progéniteurs Apicaux..... | 11 |
| 1.3.1.c Progéniteurs Basaux..... | 11 |
| 1.3.2 Voies moléculaires de la différenciation neurale | 13 |
| 1.3.2.a Pluripotence et Auto-Renouvellement | 14 |
| 1.3.2.b Neurogenèse..... | 15 |
| 1.3.2.c Astrogenèse | 16 |
| 1.3.2.d Oligodendrogenèse | 17 |
| 1.3.3 L'épigénétique dans le développement neural..... | 18 |
| 1.3.3.a Le Complexe BAF | 19 |
| 1.3.3.b Le Complexe NuRD | 22 |

| | |
|--|----|
| 1.3.3.c <i>TRRAP</i> et le complexe TIP60 | 24 |
| 1.4. Retard du développement, déficit intellectuel, et autisme | 27 |
| 1.4.1 Définitions et informations générales | 27 |
| 1.4.2 Caractérisation des phénotypes étudiés | 29 |
| 1.4.2.a Les BAFopathies et <i>SMARCC2</i> | 29 |
| 1.4.2.b Syndrome Neurodéveloppemental associé à <i>CHD3</i> | 31 |
| 1.4.2.c Syndrome Neurodéveloppemental associé à <i>TRRAP</i> | 32 |
| 1.5. Gènes à l'étude..... | 33 |
| 1.5.1 SMARCC2: SWI/SNF Related, Matrix Associated, Actin Dependent Regulator of Chromatin Subfamily C Member 2..... | 33 |
| 1.5.2 CHD3: Chromodomain Helicase DNA binding protein 3 | 35 |
| 1.5.3 TRRAP: Transformation/Transcription domain Associated Protein | 36 |
| 1.6 Problématique et Hypothèses..... | 37 |
| 2. Matériel et Méthodes | 38 |
| 2.1. Génération de lignées fibroblastiques et lymphoblastoïques et extraction d'ARN | 38 |
| 2.1.1 Génération de lignées et extraction d'ARN | 38 |
| 2.1.2 Séquençage et qPCR | 40 |
| 2.2. Analyse Bioinformatique | 42 |
| 2.2.1 Séquençage de l'ARN..... | 42 |
| 2.2.2 ATACseq | 43 |
| 2.2.3 Analyse du regroupement statistique des mutations..... | 43 |
| 3. Résultats..... | 44 |
| 3.1 Des mutations dans le gène <i>TRRAP</i> sont associées à un syndrome neurodéveloppemental et induisent des changements d'expression génique..... | 44 |
| 3.1.1 Mise en situation..... | 44 |
| 3.1.2 Article 1 | 45 |
| 3.1.2. Vérification des résultats de RNAseq..... | 85 |
| 3.1.3 Séquençage de l'ARNm de LCLs d'individus TRRAP..... | 86 |
| 3.2. Des mutations dans <i>SMARCC2</i> sont associées à un syndrome neurodéveloppemental et causent des changements d'expression génique | 88 |
| 3.2.1 Mise en situation | 88 |
| 3.2.2 Article 2 | 89 |

| | |
|--|-----|
| 3.3 Des mutations dans CHD3 sont associées à un syndrome neurodéveloppemental et induisent des changements d'expression génique..... | 131 |
| 3.4 Analyse de l'ouverture de la chromatine dans des fibroblastes d'individus TRRAP, SMARCC2, et CHD3..... | 133 |
| 4. Discussion..... | 136 |
| Bibliographie..... | 143 |
| Annexe I : Gènes différentiellement régulés dans les fibroblastes ou LCLs d'individus porteurs de variants dans <i>SMARCC2</i> , <i>TRRAP</i> , et <i>CHD3</i> | i |

Liste des tableaux

| | |
|--|-----|
| Tableau I - Syndromes associés aux hélicases de la famille CHD | 32 |
| Tableau II - Syndromes Associés aux KATs/KDACs..... | 33 |
| Tableau III - Cellules utilisées dans les expériences. | 39 |
| Tableau IV: Séquence des amorces utilisées pour la quantification par PCR en temps réel. | 42 |
| Tableau V - Gènes différentiellement régulés dans les LCLs d'individus porteurs de variants dans TRRAP impliqués dans le développement neural. | 87 |
| Tableau VI - Gènes différentiellement régulés dans des fibroblastes d'invididus porteurs de mutations dans <i>CHD3</i> impliqués dans le développement neural. | 132 |
| Tableau VII - Gènes significatifs en RNAseq et ATACseq pour des fibroblastes TRRAP. | 135 |

Liste des figures

| | |
|---|----|
| Figure 1: L'épigénétique régule l'expression génique en modifiant la chromatine (Mirabella et al, 2016)..... | 2 |
| Figure 2: Les résidus des queues d'histones peuvent être modifiées. (Lawrence et al., 2016) | 7 |
| Figure 3: Différenciation séquentielle des progéniteurs neuraux au cours du développement mammifère | 12 |
| Figure 4: Voies moléculaires impliquées dans la différenciation de cellules neurales | 13 |
| Figure 5: Les régulateurs de la chromatine sont impliqués dans le développement neural. (Sokpor et al, 2018) | 18 |
| Figure 6: Répartition des mutations associées à un syndrome neurodéveloppemental dans les gènes du complexe BAF (Adapté de Mashtalir, et al. 2018) | 30 |
| Figure 7: Localisation des mutations dans <i>CHD3</i> associées au syndrome Snijders-Blok- Campeau | 35 |
| Figure 8: Genotype-phenotype correlation associated with TRRAP variants..... | 55 |
| Figure 9: <i>TRRAP</i> sequence is intolerant to missense variants..... | 57 |
| Figure S2. Fibroblasts harboring the p.(Leu805Phe) or p.(Trp1866Cys) variants have differential gene expression patterns. | 83 |

| | |
|--|------------|
| Figure 10: Vérification des gènes sélectionnés du RNAseq par qPCR. | 86 |
| Figure 11 : Les LCLs d'individus TRRAP présentent des gènes différentiellement régulés. | 87 |
| Figure 12: Pictures of ten of the individuals with SMARCC2 variants..... | 94 |
| Figure 13: SMARCC2 (BAF170) main domains and variants. | 96 |
| Figure 14: Splicing variant analyses in Individual 8 and 4..... | 99 |
| Figure 15: Fibroblasts harboring p.Leu610Pro or p.Asn134Asp SMARCC2 mutation have differential gene expression patterns when compared to fibroblasts from healthy controls. | 102 |
| Figure 16: Les fibroblastes porteurs de mutations dans <i>CHD3</i> présentent un profil d'expression génique différent.. | 131 |

Liste des sigles

TSS : Transcription Start Site
DMR : Differentially Methylated Region
MBP : Methyl Binding Protein
CPN : Cœur Protéique Nucléosomique
NFR : Nucleosome Free Region
CRC : Complexe de remodelage de la chromatine
HAT : Histone Acetyl-Transferase
KAT : Lysine Acetyl-Transferase
HDAC : Histone Deacetylase
KDAC : Lysine Deacetylase
SNC : Système Nerveux Central
CNE : Cellule Neuroépithéliale
CSN : Cellule Souche Neurale
PA : Progéniteur Apical
ZV : Zone Ventriculaire
PARG : Progéniteur Apical Radial Glial
API : Apical Progenitor Intermediate
LCR Liquide Céphalo Rachidien
PB : Progéniteur Basal
ZSV : Zone Sub-Ventriculaire
PNE : Progéniteur Neuronale
PAC : Progéniteur Astrocytaire
POC : Progéniteur Oligodendrocytaire
TCF : T-Cell Factor
LEF : Lymphocyte Enhance Factor
aPKC : Protéine Kinase C atypique
SHH : Sonic Hedgehog
PTCH : Patched
SMO : Smoothened
NICD : NOTCH Intracellular Domain

bHLH : basic Helix Loop Helix
 BMP : Bone Morphogenic Protein
 TGFb : Transformation Growth Factor beta
 CNTF : Ciliary Neurotrophic Factor
 LIF : Leukemia Inhibitory Factor
 GFAP : Glial Fibrillary Acidic Protein
 ISWI : Imitation Switch
 CHD/NuRD : Chromodomain Helicase DNA binding / Nucleosome Remodeling Deacetylase
 BAF : Brg or Brm Associated Factors
 SWI/SNF : Switch / Sucrose Non Fermentable
 esBAF : Embryonic Stem BAF
 npBAF/nBAF : neural progenitor / neuronal BAF
 PRC1 : Polycomb Repressor Complex 1
 PKCd : Protein Kinase C delta
 iPSC : induced Pluripotent Stem Cell
 ATM : Ataxia telangiectasia mutated
 NDD : NeuroDevelopmental Disorder
 DD/RDG : Délai du Développement / Retard du Développement Global
 DI : Déficit Intellectuel
 TSA : Trouble du Spectre de l'Autisme
 VTP : Variant Tronquant la Protéine
 pLI : Probability of Loss-of-Function Intolerance
 TDISS : Troubles de Déficience Intellectuelle liés au SWI/SNF
 CSS : Coffin Siris Syndrome
 NBS : Nicolas-Barais Syndrome
 PHD : Plant Homeodomain
 Chromo : Chromatin Organization Modifier
 DUF : Domain of Unknown Function
 ATACseq : Assay for Transposase Accessible Chromatin Sequencing
 RNAseq : RNA Sequencing
 PBS : Phosphate Buffer Saline
 FBS : Fetal Bovine Serum

DMEM : Dulbecco's Minimum Essential Media

RPMT : Roswell Park Memorial Institute

LCL : Lymphoblastoic Cell Line

GO : Gene Ontology

DEG : Differentially Expressed Gene

Remerciements

Je tiens tout d'abord à remercier Philippe pour m'avoir donné une chance malgré mon dossier académique peu performant; pour son soutien continu et de m'avoir laissé développer mes capacités en bio-informatique en dépit de mes connaissances initiales presque nulles dans le domaine. Je remercie particulièrement Justine, qui a été un pilier de mon cheminement académique et personnel, une personne inspirante qui m'a permis de rester saine toutes ces années. Je remercie Nissan, Mai, les stagiaires et employés du laboratoire qui m'ont beaucoup appris et ont su garder une bonne ambiance. Je remercie aussi mes amis et ma famille qui m'ont soutenue et aidée.

Introduction

1.1 Remarques générales

Depuis la découverte de la double hélice d'ADN par Watson, Crick et Franklin¹, et de la structure des acides nucléiques par Pauling et Corey en 1953², le dogme principal de la biologie moléculaire se résume en trois mots: ADN, ARN, protéine. Depuis, des avancées majeures ont permis la découverte des mécanismes de régulation opérant à travers ce dogme pour permettre d'affiner les subtilités de l'homéostasie cellulaire. Notamment, l'équilibre de l'expression d'ARN a été démontré comme étant un phénomène complexe et indispensable, reposant sur une organisation infinitésimale de régulateurs de l'expression génique. De ces régulateurs, certains agissent sur la structure même de la chromatine afin de modifier à la fois sa composition chimique au niveau de l'acide nucléique ou de l'histone, et sa compaction à l'échelle locale. Ces trois mécanismes, soit la modification directe de l'ADN, l'ajout ou la rétraction de groupes chimiques sur les octamères d'histones, et le remodelage de la chromatine forment l'épigénétique: l'information «au-dessus» de la génétique (Figure 1). Sans changer la nature même de l'ADN, l'épigénétique est, à travers l'action des régulateurs de la chromatine et par sa nature à la fois transmissible et malléable, le chef d'orchestre du destin cellulaire qui différencie un neurone d'un ostéoblaste. De par la complexité de cette organisation et le nombre de régulateurs responsables du maintien ou du changement des marques de l'épigénétique, le système est sensible et vulnérable à des variations. Ainsi, de nombreuses mutations dans ces régulateurs épigénétiques ont été rapportées comme étant pathogéniques et des causes de maladies telles que le cancer ou des troubles neurodéveloppementaux (Neurodevelopmental Disorders, ou NDDs). Dans le cadre de ce présent mémoire, des mutations dans des régulateurs épigénétiques qui n'ont précédemment pas été associés à des maladies humaines impliquant le neurodéveloppement seront étudiées afin d'établir de nouveaux liens entre régulateurs épigénétiques et le neurodéveloppement.

1.2 Les mécanismes de l'épigénétique

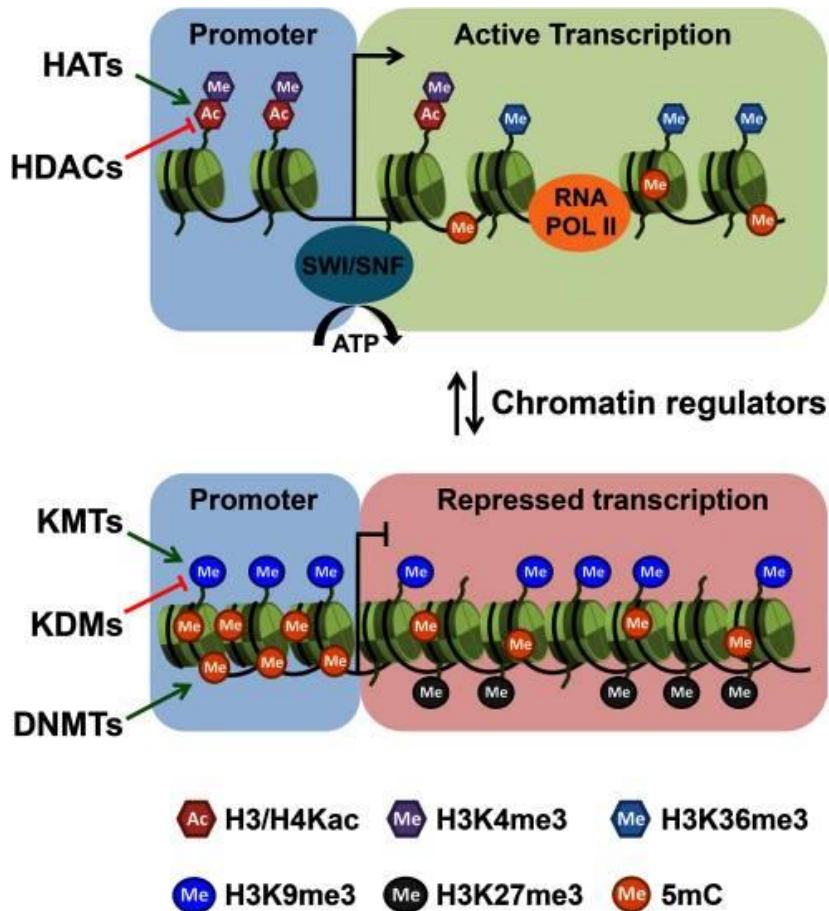


Figure 1: L'épigénétique régule l'expression génique en modifiant la chromatine (Mirabella et al, 2016). Schéma montrant les modifications possibles de la chromatine, soit les modifications d'histones (méthylation, acétylation), la modification de l'ADN (méthylation), et le déplacement de nucléosomes. La polymérase à ARN (RNA POL II) n'a accès à la chromatine que quand sa structure est accessible. HAT : Histone Acétyltransférase, HDAC : Histone Deacetylase, KMT : Lysine Méthyltransférase, KDM : Lysine Déméthylase, DNMT : Déméthylase de l'ADN.

1.2.1 L'ADN et sa composition

La molécule d'ADN, ou acide désoxyribonucléique, est porteuse de l'information génétique par sa séquence. Elle est formée d'une succession de nucléotides qui sont rajoutés les uns après les autres pour former des chromosomes allant de 250 millions de paires de bases à 46 millions pour une totalité de plus de trois milliards de paires de bases^{3,4}. Les nucléotides traditionnels sont l'adénosine (A), la Thymine (T), la Guanine (G), et la Cytosine (C). Ce dernier

nucléotide est unique dans sa capacité à être modifié de façon covalente par l'addition d'un groupement méthyle (CH_3) pour devenir du 5-méthylcytosine. Cette réaction est catalysée par des enzymes méthyltransférases à ADN (DNA methyltransferase, DNMT)⁵, tandis que la déméthylation peut se faire de façon passive pendant la réplication où l'ADN nouvellement synthétisé ne porte pas de modification et elle n'est pas rajoutée de façon subséquente⁵. Ainsi, la marque épigénétique à ce locus précis est perdue après plusieurs cycles de réplication. Sinon, la déméthylation peut se faire de façon active en plusieurs étapes qui permettront d'enlever le groupement méthyle, en substituant le nucléotide modifié par une cytosine à travers la modification par les enzymes de la famille TET, ou en excisant une région entière et la réparant avec des nucléotides non-modifiés⁶.

La méthylation de l'ADN est généralement considérée comme une marque de répression de l'expression, principalement à travers la méthylation d'îlots CpG (une Cytosine reliée à une Guanine par un seul phosphate) qui se trouvent au niveau des sites de début de transcription (transcription start-site; TSS)⁵. Près de 70% des promoteurs annotés dans le génome humain sont associés à ces régions caractérisées par un haut taux de G/C et un enrichissement de CpG, et qui permettent une régulation de l'initiation de la transcription⁵. La méthylation permet d'établir une répression à long terme d'une région transcrite car la déméthylation spontanée est un événement extrêmement rare⁶.

Chez les mammifères, les cellules embryonnaires sont dénudées d'une grande partie des marques de méthylation dans leurs premiers instants^{5,7}. Ensuite, de nouvelles marques sont établies et maintenues lors de la différenciation cellulaire, notamment pour empêcher l'expression de gènes de pluripotence, mais aussi l'expression de transposons et d'un allèle des gènes soumis à empreinte parentale^{5,8,9}. La méthylation *de novo* de l'ADN se fait sous l'action de la protéine DNMT1 tandis que la maintenance se fait par DNMT3A et B⁹. Alors, la méthylation de certaines régions du génome permet le recrutement de protéines les reconnaissant spécifiquement: les MBPs (Methyl Binding Proteins, ou Protéines de Liaison au Méthyle)^{10,11}. Les MBPs peuvent coordonner un effet répresseur en recrutant des co-répresseurs qui empêcheront la liaison de la machinerie de transcription ou d'enzymes permettant d'inverser la réaction et de retrouver un nucléotide non-modifié. Elles peuvent aussi recruter des complexes

capables de remodeler les nucléosomes l'entourant et compacter l'ADN afin de stabiliser la région réprimée^{12,13}.

1.2.2 La chromatine et les nucléosomes

Le brin d'ADN est formé de nucléotides qui forment une double hélice. Cette dernière s'enroule autour de complexes protéiques nommés nucléosomes qui permettent une compaction de la molécule d'ADN afin d'en réduire son volume. Le nucléosome comprend le cœur protéique nucléosomique (CPN) et l'ADN qui l'entoure. Le CPN est lui-même composé de sous-unités nommées histones, soit H2A, H2B, H3, et H4, qui sont présentes chacune en deux exemplaires pour former un octamère^{13,14}. Chacune de ces sous-unités possède aussi des variantes (sauf H4) codées par des gènes différents et pour lesquels différentes fonctions sont associées¹⁵. Les nucléosomes sont aussi associés à des histones de liaison H1 qui permettent de lier un CPN à l'ADN de liaison entre deux nucléosomes. Par ce fait, la région est compactée et le nucléosome est stabilisé. Les histones étant porteuses de charges positives, il est convenu que la force d'interaction entre le nucléosome et l'ADN chargé négativement se fait par des interactions électrostatiques. Un nucléosome est entouré de 145-147 nucléotides, et un nucléosome est généralement trouvé toutes les 200 paires de bases¹⁶.

La chromatine existe principalement sous deux états: l'euchromatine qui est considérée comme une conformation «ouverte» où la machinerie de transcription et d'autres complexes sont en mesure d'accéder à la chromatine, et l'hétérochromatine où la compaction de la chromatine ne permet pas la liaison de facteurs de transcriptions et d'autres complexes pro-transcriptionnels¹⁷⁻¹⁹. Ces régions sont limitées dans l'espace, c'est à dire qu'à un moment donné, certains domaines de la chromatine se trouvent sous forme d'euchromatine tandis que d'autres sont maintenus fermés. De plus, l'hétérochromatine peut être fermée en tout temps, comme dans les régions des gènes soumis à l'empreinte; elle est nommée hétérochromatine constitutive^{9,18}. Sinon, elle peut être fermée uniquement à certains moments selon le type cellulaire, le stade de différenciation, ou même la phase du cycle cellulaire²⁰⁻²³. Tous ces mécanismes sont dépendants des nucléosomes et de leur position dans la chromatine, où le consensus général tend vers une accumulation de nucléosomes dans des régions fermées du

génomique, à l'inverse de régions relativement dépourvues de nucléosomes (Nucleosome Free Regions, NFRs) lorsqu'elles sont transcriptionnellement actives²⁴.

Les complexes responsables du remodelage de la chromatine sont multiples et permettent plusieurs réactions qui modifieront la position ou la composition du CPN. Ils peuvent, par leur fonction hélicase ATP-dépendante, déplacer des nucléosomes à la paire de base près le long d'un brin d'ADN^{25,26}. De plus, ils peuvent modifier un nucléosome en expulsant une ou plusieurs histones, déstabilisant le nucléosome²⁷. La composition unique du complexe déterminera les séquences cibles en combinant des protéines ayant des domaines de reconnaissance de séquences ou de marques épigénétiques spécifiques.

1.2.3 Les histones et leurs variantes

Le CPN est formé de deux copies de chaque histone du cœur, et chacune d'entre elle peut exister sous leur forme canonique, ou sous la forme d'une variante. Les variantes sont codées par des gènes différents et sont aussi stabilisées dans le noyau par des chaperonnes différentes des histones constitutives¹⁵. Du fait de leur propension à sédimenter ensemble, il est indispensable de les associer à une chaperonne qui pourra aussi leur donner une spécificité de ciblage dans le génome par ses sites de liaison distincts. Chaque chaperonne est porteuse d'un dimère d'histones H3/H4 ou H2A/H2B, où chacune des histones peut se trouver sous sa forme canonique ou d'une variante, de sorte à créer des homo-dimères ou hétéro-dimères^{15,27}. Contrairement aux histones canoniques, la transcription des gènes codants pour les variantes n'est pas dépendante de la réplication. Au contraire, leur transcription est régulée par différents signaux, tels que la réparation aux dommages à l'ADN, l'augmentation de la transcription, ou la différenciation cellulaire^{15,27-29}. L'échange d'histones peut se faire par le désassemblage total de nucléosomes par des complexes de remodelage de la chromatine, puis la reformation de nucléosomes en utilisant de nouveaux dimères d'histones pouvant comprendre des variantes. Sinon, un nucléosome peut aussi être déstabilisé et un unique dimère déplacé du CPN. Ensuite, un nouveau dimère est déposé par une chaperonne et restabilise la structure du nucléosome^{28,30}.

Le dépôt d'histones ayant des séquences légèrement différentes, et donc des propriétés chimiques et physiques différentes permet à plusieurs mécanismes épigénétiques de se produire. Premièrement, des protéines comprenant des domaines de liaison à ces variantes d'histones

peuvent s'y lier, et donc conférer une spécificité de liaison à la chromatine qui va au-delà de la simple séquence. De ce fait, l'histone H2A.X est spécifiquement phosphorylée (γ H2AX) aux sites de bris de l'ADN, permettant le recrutement et le maintien à ces sites, ou *foci*, de la machinerie de réparation des dommages à l'ADN^{31,32}. Notamment, la chromatine est décondensée au niveau des *foci* de dommages à l'ADN, probablement par les complexes de remodelage de la chromatine qui y sont recrutés par affinité pour γ H2AX³². Ensuite, les propriétés de différentes variantes confèrent au nucléosome une stabilité en conséquence. Certaines variantes étant plus sensibles *in vitro* à un gradient de sel; elles se délogent plus facilement de la chromatine *in vitro*, probablement à cause d'interactions moins fortes avec le reste du CPN ou l'ADN^{15,30}. Ainsi, la variante H3.3 est plus présente au niveau des TSS, correspondant au besoin de la polymérase à ARN POL II de déplacer les nucléosomes lors de la transcription^{30,31}.

1.2.4 Les modifications d'histones

Chaque histone forme une structure particulière incluant une queue N-terminale et un cœur globulaire positionnée plus centralement au nucléosome. Sur ceux-ci, certains acides-aminés peuvent être modifiés chimiquement pour rajouter de façon réversible des groupements permettant d'altérer les interactions entre nucléosome et ADN, ou de changer les interactions spécifiques avec d'autres protéines (Figure 2)³³. Toutes les histones peuvent être modifiées, bien que les histones du CPN soient les plus étudiées^{16,34}. Les queues des histones étant positionnées à l'extérieur du CPN, elles sont donc plus accessibles et plus facilement modifiables, mais des modifications du cœur d'histone sont aussi possibles. Malgré leur rôle important dans la biologie cellulaire, les queues d'histones peuvent être complètement tronquées sans impact sur la stabilité du CPN³⁵. C'est donc leur rôle dans la formation de la structure locale et globale de la chromatine qui est indispensable.

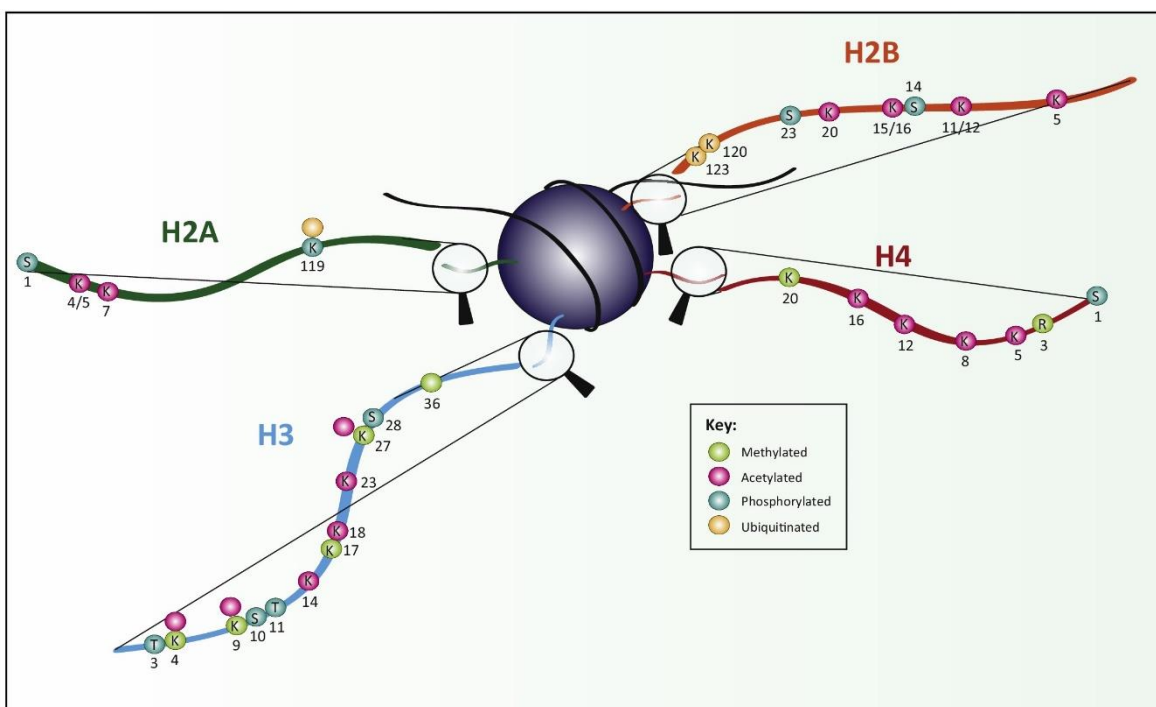


Figure 2: Les résidus des queues d'histones peuvent être modifiés. (Lawrence et al., 2016). Chaque résidu est indiqué par son code uni-lettre, soit T : Thréonine, S : Sérine, K : Lysine, R : Arginine. La modification du résidu est indiquée par sa couleur, soit vert : méthylation, rose : acétylation, bleu : phosphorylation, jaune : ubiquitination.

Ainsi, sachant que les interactions se font principalement grâce aux charges positives des histones en complément des charges négatives de l'ADN, l'ajout de groupements acétyles permet la neutralisation de l'histone et donc la déstabilisation de ces interactions^{16,33}. Par conséquent, la structure de la chromatine décondensée et plus accessible pour la transcription. L'ajout de groupements acétyles sur les lysines d'histones est médié par des Histones Acétyl-Transférases (HAT) ou Lysine Acétyl-Transférases (KAT), tandis que le retrait de ces modifications se fait par des Histones Déacétylases (HDAC) ou Lysine Déacétylases (KDAC)¹⁶. Les HAT et HDAC sont spécifiques aux histones, tandis que les KAT et KDAC peuvent aussi acétyler d'autres protéines. Chacune permet de modifier une ou plusieurs lysines distinctes, permettant l'obtention de profils uniques selon l'expression de certaines HAT ou HDAC. Certaines acétylations sont connues comme étant des signes évidents d'une région transcriptionnellement active, comme H3K27Ac, H3K4Ac, ou encore H2AK7Ac^{16,36}.

La seconde modification d'histone est la méthylation de résidus lysines. Contrairement à l'acétylation, la méthylation ne cause pas de modification au niveau de la chimie du nucléosome¹⁶. Par contre, elle permet le recrutement d'autres facteurs qui amplifient le signal à travers la poly-méthylation (jusqu'à trois groupements sur la même lysine) et le recrutement de complexes de remodelage de la chromatine comme PRC1/2³⁷, ou même empêche le recrutement de certains facteurs en cachant le site de liaison. De plus, il n'y a pas de consensus quant au rôle de la méthylation des lysines d'histones. En effet, elle a généralement un rôle de répression, comme pour H3K27me3 qui est notamment associée à la répression par son rôle antagoniste à l'acétylation, et est une marque indispensable pour l'inactivation du chromosome X³⁸. Certaines méthylations sont aussi des marques de l'hétérochromatine constitutive comme H4K20me3, H3K9me3, et H3K79me3. Par contre, H3K36me3 est une marque enrichie dans les régions transcriptionnellement actives, mais aussi présente dans l'hétérochromatine constitutive de certains gènes soumis à l'empreinte^{39,40}. Enfin, H3K4me2 est une marque épigénétique intéressante car elle est présente en complément avec des marques de répression comme H3K27me3 dans des régions qui ne sont pas transcriptionnellement actives, mais qui le deviendront dans le développement, acquérant à la place la marque activatrice H3K27ac^{41,42}. Ce phénomène forme des régions de préparation à la transcription (« poise ») qu'on appelle bivalentes, c'est à dire de l'hétérochromatine facultative. Autre que les lysines, certains résidus

d'arginine peuvent aussi être méthylés, et ces méthylations sont principalement associées à l'activation de la transcription^{16,36}.

Les queues d'histones étant à l'extérieur du CPN, elles sont bien plus accessibles par les enzymes capables de les modifier. Pourtant, un nombre de résidus localisés sur la surface globulaire du nucléosome ont été détectés portant des modifications, et sont nécessaires pour plusieurs mécanismes cellulaires. Par exemple, H3K79me3 est indispensable pour le raccourcissement des télomères causé par la sénescence⁴³, et H3K122Ac augmente l'éviction de nucléosomes *in vitro* et stimule la transcription en s'intercalant directement entre le CPN et le brin d'ADN qui l'entoure^{16,44}.

D'autres modifications peuvent être sources de signaux sur les nucléosomes, comme la phosphorylation, l'ubiquitination, l'ADP ribosylation, la déamination, ou la sumoylation¹⁵. Notamment, la phosphorylation par les kinases Aurora A et B sont nécessaires lors de la mitose pour l'hypercondensation des chromosomes et le recrutement d'enzymes capables de déacétyler les histones du CPN pour stimuler des interactions inter-nucléosomes^{45,46}.

1.3. Le développement neural

1.3.1 Différenciation séquentielle des différents types cellulaires du cerveau

1.3.1.a Cellules Souches Neurales et Cellules Neuroépithéliales

Le système nerveux central (SNC) est formé de cellules provenant de l'ectoderme qui se différencient en les premières cellules neurales: les cellules neuroépithéliales (CNE). Celles-ci donneront lieu aux cellules souches neurales (CSNs), ou cellules progénitrices neurales, la source de toutes les cellules neurales du cerveau^{47,48}. Sachant que la totalité des cellules neurales proviennent d'une petite population de cellules, et que le destin cellulaire est engagé relativement tôt dans la différenciation de la cellule, il est important de différencier les deux types de divisions qu'elles peuvent engager. La première est la division dite proliférative, ou symétrique, tandis que la seconde est de différenciation, ou asymétrique. Le caractère symétrique ou asymétrique est coordonné par l'orientation du fuseau mitotique en comparaison à la composition de la membrane de la cellule. En effet, certains éléments du cortex cellulaire

et du cytoplasme peuvent être ségrégés à certaines régions de façon parallèle ou perpendiculaire au fuseau mitotique^{48,49}. Si la séparation est parallèle au fuseau, alors la cytokinèse permet la séparation de deux cytoplasmes et deux membranes porteurs de facteurs différents. Par conséquent, il est possible d'obtenir deux cellules filles différentes dans une même division cellulaire. Ainsi, dans le cas d'une division asymétrique d'une CSN, une des cellules filles conserve ses traits de cellule progénitrice, tandis que l'autre peut engager un destin cellulaire terminal. De plus, les divisions symétriques peuvent se produire soit pour amplifier la population de progéniteurs souches qui n'ont pas enclenché un profil de différenciation, ou des progéniteurs qui sont déjà destinés à un type cellulaire particulier afin de contrôler spécifiquement le nombre de cellules correspondant à chaque catégorie⁴⁷⁻⁴⁹. En général, le terme de CSN est controversé car une cellule neurale est déjà partiellement différenciée mais sera utilisée pour caractériser les cellules donnant lieu aux trois types majoritaires de cellules du cerveau: les neurones, les astrocytes, et les oligodendrocytes.

Les cellules NE ainsi que les autres progéniteurs neuraux non-souches (qui ne peuvent pas se différencier en tous les types neuraux, mais ne sont pas en phase terminale de différenciation) sont capables de se diviser de façon asymétrique et symétrique, bien que la division symétrique de progéniteurs non-souches déjà enclenchés dans un sous-type neural soit généralement courte (2-3 divisions)⁴⁸. On parle de cellules souches neurales, ou cellules progénitrices neurales, pour des CNEs qui ont acquis des marqueurs de différenciation, mais qui sont encore multipotentes. La grande majorité des cellules neurales progénitrices est épuisée lors du développement, mais certaines cellules appelées cellules souches adultes sont conservées et peuvent être ré-activées plus tard, bien que leurs fonctions et mécanismes soient peu connus^{50,51}. Les CSNs sont, à tous les stades du développement, organisées parallèlement les unes aux autres, avec un côté orienté vers l'intérieur, que l'on nomme la surface apicale; et un côté orienté vers l'extérieur, soit la surface basale. Quand il y a division, les cellules différenciées migrent vers l'extrémité basale de sorte que les cellules les plus jeunes soient à l'extrémité basale du cerveau. Il y a donc un gradient de cellules formées au cours du temps de la surface apicale à la surface basale, formant des couches successives de cellules spécifiques. C'est la continuité des CNEs et des premiers progéniteurs entre les deux extrémités qui permet

de guider les cellules différenciées vers la lamina basale en leur permettant de migrer le long des projections des progéniteurs.

1.3.1.b Progéniteurs Apicaux

La première grande catégorie de CSNs est les progéniteurs apicaux (PAs). Ces derniers sont différents des CNEs par l'acquisition de marqueurs astrogliaux tels que GLAST ou BLBP qui permettent de les différencier des progéniteurs du système nerveux périphérique et des progéniteurs de la rétine⁴⁹. Chaque PA est généralement restreint dans sa capacité à générer un type cellulaire selon le type de marqueur qu'il exprime, qui lui même dépend du stade de développement. En effet, les PAs acquièrent successivement différents marqueurs leur permettant de se diviser en un seul type cellulaire, principalement neurone, astrocyte, puis oligodendrocyte^{52,53}.

Le noyau des PAs migre de façon dépendante du cycle cellulaire le long de l'axe apico-basal. Ainsi, en stade G1, leur noyau se trouve accolé à la surface apicale puis migre radialement vers la lamina basale pour la réplication. Chez les PAs, contrairement aux CNEs, le noyau n'atteint pas l'extrémité basale, et est restreint dans la couche appelée zone ventriculaire (ZV), où il n'y a que des progéniteurs. Ensuite, le noyau rebrousse chemin vers l'extrémité apicale où se produit sa mitose. S'il y a formation d'une cellule différenciée, elle migre au dessus des noyaux des PAs hors de la ZV^{52,54}. La majorité des PAs sont des PA radiaux gliaux qui conservent un processus apical et basal. Certains progéniteurs apicaux sont nommés progéniteurs apicaux intermédiaires (APIs) puisqu'ils se différencient des PA radiaux gliaux par la rétraction de leur processus basal pour la mitose, les forçant dans une division symétrique neurogénique⁵⁴. La membrane apicale des PAs ne représente que 1 à 2% de la membrane totale, mais contient le cilium primaire qui s'étend vers la lumière du ventricule du cortex en développement⁴⁸. Par cette connection, les cellules apicales maintiennent un contact avec le Liquide Céphalo-Rachidien et sert d'antenne pour recevoir les signaux moléculaires en provenant.

1.3.1.c Progéniteurs Basaux

La seconde classe de progéniteur neural est les progéniteurs basaux (PBs). Ces progéniteurs sont formés par la division de PAs, mais ce sont toujours des cellules capables de

proliférer et de donner lieu à des types cellulaires différents^{47,48,55}. Les PBs sont différents des PAs par leur mouvement lors du cycle cellulaire: à la première division d'un PB généré depuis un PA, son noyau migre vers la lamina basale, et la cellule se détache de la surface apicale de sorte à perdre le contact avec les signaux apicaux^{48,53}. Alors, les PBs sont ensuite générés, en différents sous-types. Contrairement aux PAs, la mitose des PBs a lieu à l'extrémité basale de la ZV, et leur noyau migre très peu dès qu'il y est localisé. Tous les PBs se divisent peu, soit une à trois divisions, et de façon uniquement symétrique, premièrement proliférative puis neurogénique⁵⁴. Plusieurs types de PBs existent selon leur morphologie, catégorisés en deux classes majeures: les PBs radiaux et les PBs intermédiaires. Seuls les PB radiaux maintiennent un processus basal, les PB intermédiaires se trouvant dans la ZV sans attache aux extrémités^{48,53}.

Les PBs forment au cours du développement la zone sub-ventriculaire (ZSV) qui se trouve entre la ZV et la surface basale. Celle-ci est suggérée comme étant responsable de la croissance de la quantité de neurones au cours de l'évolution. En effet, un lien entre la taille de la ZSV et l'avancement dans l'arbre phylogénétique a été observé, indiquant qu'un plus grand nombre de progéniteurs donne lieu à un plus grand cortex; et donc une plus grande complexité neurale⁵⁶.

Après une différenciation terminale, les cellules formées sont des précurseurs des types cellulaires majoritaires dans le cortex: des précurseurs neuronaux (PNE), astrocytaires (PAC), ou oligodendrocytaires (POC) (Figure 3). Chacun opère ensuite à plusieurs étapes de maturation afin de former des cellules opérationnelles.

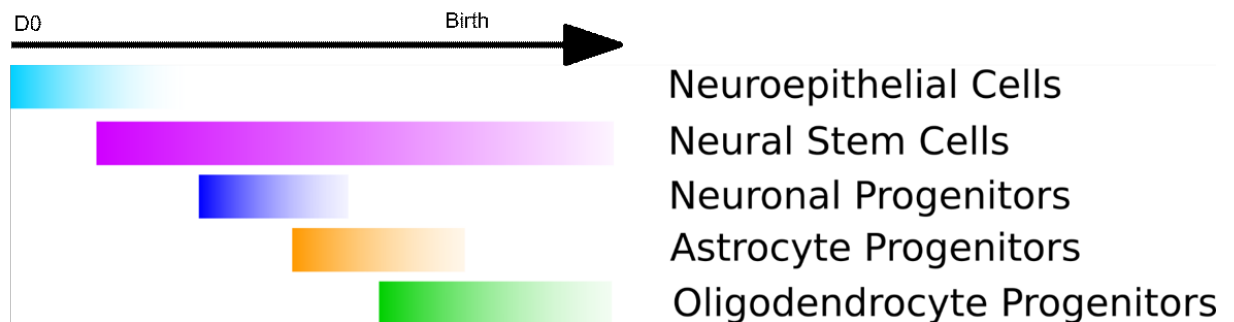


Figure 3: Différenciation séquentielle des progéniteurs neuraux au cours du développement mammifère. Schéma démontrant la progression de la formation de chacun des progéniteurs des cellules du cortex cérébral dans le développement des mammifères. Les cellules neuroépithéliales acquièrent rapidement les marqueurs des cellules souches neurales qui peuvent former séquentiellement en progéniteur neuronal,

progéniteur astrocytaire, et enfin progéniteur oligodendrocytaire selon la phase du développement à laquelle la division prend lieu.

1.3.2 Voies moléculaires de la différenciation neurale

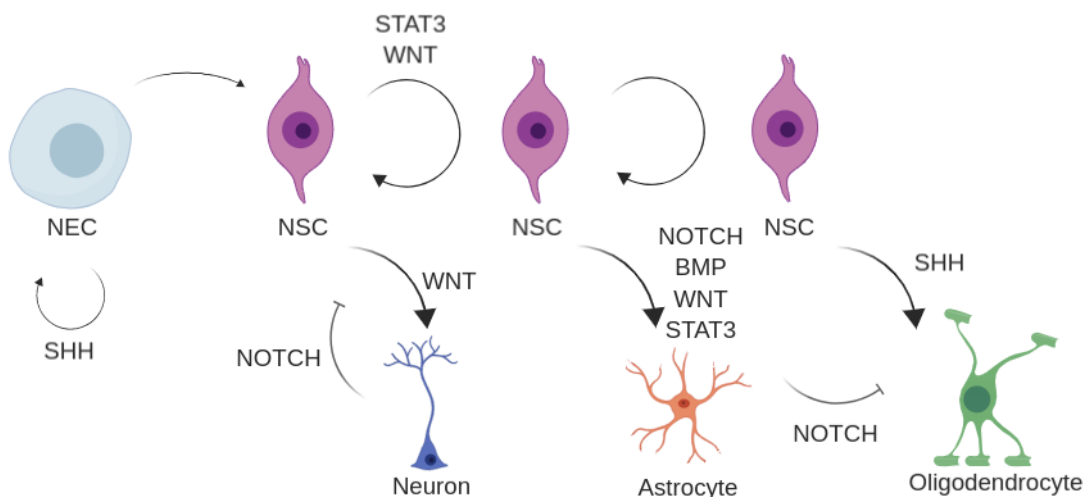


Figure 4: Voies moléculaires impliquées dans la différenciation de cellules neurales. L'activation ou la répression de voies moléculaires est indispensable pour la maintenance de l'état pluripotent ou la différenciation en progéniteur spécialisé. NEC : Cellule Neuroépithéliale, NSC : Cellule Souche Neurale. Les flèches circulaires indiquent l'auto-renouvellement.

Chaque étape du développement des CNEs jusqu'aux cellules différenciées est contrôlé par des changements spécifiques des gènes exprimés dans les cellules progénitrices (Figure 4). Ainsi, plusieurs voies biologiques sont responsables de la formation de certains types cellulaires, par leur présence ou leur absence. De plus, l'activation d'une même voie moléculaire dans des contextes différents par le temps ou l'espace peut avoir une conséquence radicalement différente. Deux mécanismes principaux agissent de façon antagoniste afin de déterminer le type de division que les progéniteurs réalisent: le maintien de la multipotence et la différenciation. Ainsi, la multipotence insinue l'expression de facteurs appartenant aux cellules « souches » ou prolifératives, et la différenciation signifie la répression de ceux-ci, et l'expression de facteurs spécifiques à un type cellulaire. De ce fait, ce n'est pas qu'en réprimant les facteurs de prolifération qu'un progéniteur peut donner lieu à une cellule différenciée, mais aussi en induisant une spécification. De plus, une hypothèse majoritaire de la sélection de la voie de

différentiation ou de l'auto-renouvellement est que la longueur du cycle cellulaire serait un des ses facteurs déterminants^{57,58}. En effet, un cycle cellulaire court ne serait pas propice à l'accumulation de facteurs de polarité et promouvrait une différenciation symétrique. Au contraire, le cycle cellulaire d'une cellule progénitrice est significativement ralenti quand elle va procéder à une division neurogénique⁵⁹, et l'inhibition temporaire de l'avancement du cycle cellulaire est suffisante pour induire une division neurogénique⁶⁰.

1.3.2.a Pluripotence et Auto-Renouvellement

En général, le développement des cellules du SNC se fait de façon séquentielle, soit en premier la prolifération des CSNs, puis l'acquisition de marqueurs gliaux leur permet de se différencier. Les premières cellules à être formées sont des PNE, puis les CSNs gagnent la capacité de se différencier alternativement en PAC, et enfin la capacité de former des POC. L'activation successive de différentes combinaisons de voies biologiques permet la spécification de chacun des types. Dans les CNEs, l'activation de la voie WNT/ β -caténine est nécessaire pour la maintenance de l'état pluripotent tôt dans leur développement⁶¹⁻⁶³. Cette voie biologique est activée de façon canonique par la liaison de protéines de la famille WNT à un récepteur extracellulaire de la famille Frizzled. La voie d'activation canonique implique la stabilisation de la β -caténine en inhibant sa dégradation par GSK-3 β , lui permettant de transloquer vers le noyau et s'associer à des facteurs de transcription TCFs (T-cell factors, facteurs des cellules T) ou LEFs (Lymphocyte Enhance Factor, Facteur de renforcement des lymphocytes)^{64,65}. En effet, la β -caténine est nécessaire pour maintenir la polarité apico-basale des CNEs en induisant l'expression de la Protein Kinase C atypique (aPKC) grâce à l'activation de la transcription dépendante de TCF⁶⁶. Aussi, la β -Caténine serait responsable de la formation d'une surface apicale suffisamment grande pour que les deux cellules filles soient porteuses de suffisamment des facteurs de multipotence, quoi que soit l'axe de division. La voie WNT/ β -Caténine semblerait promouvoir la prolifération de CNEs en induisant les cyclines D1 et D2, ainsi que c-Myc, ce qui semble suivre l'hypothèse du cycle cellulaire contrôlant l'auto-renouvellement^{62,64,66}. Dans la rétine, la voie WNT/ β -Caténine induit l'expression de *SOX2*, un facteur de transcription indispensable pour l'auto-renouvellement et qui inhibe la différenciation^{62,66}. Par contre, *SOX2* est responsable d'une boucle de rétroaction négative où il réprime la voie WNT pour permettre aux CSN de se différencier⁶⁶.

La prolifération des CNEs est particulièrement dépendante de la voie Sonic Hedgehog (SHH). Sa voie canonique implique la liaison du ligand SHH à son récepteur extracellulaire Patched (PTCH), induisant la relâche de la répression de Smoothened (SMO), ce qui entraîne l'activation des facteurs de la famille GLI⁶⁷. Ces facteurs peuvent être activateurs ou répresseurs selon le contexte, mais il est connu que *SHH* est nécessaire pour la prolifération des CNEs *in vivo*, et que l'ajout de SHH ectopique *in vivo* induit le doublement de la prolifération de CSNs provenant du tube neural, ainsi que la prolifération de progéniteurs *in vitro*⁶⁷⁻⁶⁹. Aussi, la voie SHH inhibe la différenciation des CSN *in vivo*^{67,69}.

La voie moléculaire NOTCH dépend d'une interaction entre deux cellules, où la première exprime un ligand de la voie NOTCH, tandis que la seconde y exprime un récepteur spécifique. L'interaction induit le clivage du récepteur NOTCH pour produire le Domaine NOTCH Intracellulaire (NICD, NOTCH Intracellular Domain) qui transloque vers le noyau et s'associe avec la protéine CSL (CBF1 ou RBPJ) pour l'activer et former un complexe activateur de la transcription⁷⁰. La fonction de la voie NOTCH est particulière: l'engagement dans une voie neuronale active l'expression de gènes proneuraux, tels que *ASCL1* ou *NEUROG1*, qui eux-même induisent l'expression de ligands de NOTCH à la surface des cellules en différenciation⁷⁰⁻⁷². De ce fait, la voie NOTCH est activée dans les cellules avoisinantes, permettant un ralentissement dans la réponse à des signaux extracellulaires, ce qui inhiberait la différenciation^{70,71}. Dans ces cellules là, des facteurs de transcription bHLH (basic Helix Loop Helix, Hélice Boucle Hélice basique) comme *HES1* et *HES5* sont exprimés en aval de la voie NOTCH, ce qui inhibe l'expression de régulateurs de la différenciation comme *SOX21* et de gènes neurogéniques comme *ASCL1*^{72,73}.

1.3.2.b Neurogenèse

Au cours du développement neural, la voie WNT/ β -Caténine s'éloigne de son pouvoir prolifératif pour induire à la place l'arrêt du cycle cellulaire et la différenciation neuronale⁶¹. L'accumulation nucléaire de β -Caténine est suffisante pour induire la différenciation neuronale⁷⁴⁻⁷⁶. De plus, la voie β -Caténine active des facteurs neurogéniques comme *NEUROG1* et *NEUROG2*, des facteurs de transcriptions indispensables pour l'engagement vers un destin neuronal^{61,76}.

Les BMPs (Bone Morphogenic Proteins, Protéines Morphogènes de l'Os) font partie de la famille des TGF β (Transforming Growth Factor β , Facteur de Croissance Transformateur β). Ils sont sécrétés et activent des récepteurs à la surface des cellules BMPRI et BMPRII, des récepteurs tyrosine kinase (RTK)⁷⁷. La transduction du signal se fait ensuite par la phosphorylation de protéines SMAD qui peuvent ensuite former des complexes avec la protéine SMAD4, ou co-SMAD, et transloquer vers le noyau pour activer des gènes en se liant à la séquence de réponse spécifique aux BMPs dans leur promoteur⁷⁷. Dans les premières CSNs, les BMPs activent spécifiquement BMPRIa, qui promeut la prolifération mais qui induit aussi l'expression de *BMPRIb*^{61,74,77,78}. C'est ensuite l'atteinte d'un seuil spécifique de BMPRIb qui change l'activité des BMPs pour activer l'expression de gènes neurogéniques et induire l'arrêt du cycle cellulaire^{61,77}.

1.3.2.c Astrogenèse

Les cellules non-neuronales du cerveau sont principalement des astrocytes et des oligodendrocytes. Afin de produire ces cellules gliales, il est nécessaire pour les CSNs de réduire l'expression des gènes neurogéniques qu'elles exprimaient afin de produire des PNEs par division asymétrique, et d'exprimer en leur place des marqueurs gliaux. Même en culture, les CSNs produisent premièrement des PNEs, puis des PACs, et enfin des POCs, signifiant qu'il existe des mécanismes intrinsèques aux CSNs qui régulent la formation séquentielle de chacun des types cellulaires⁷⁹. En effet, les neurones premièrement formés sont la source de l'astrogenèse à travers leur expression de CTF1, un gène indispensable pour la formation d'astrocytes, et qui permet d'instruire les CSNs parentes de procéder au changement vers la gliogenèse⁷⁹⁻⁸¹. CT-1, la cytokine de la famille de l'interleukine-6 codée par le gène CTF1, active la voie JAK/STAT dans les CSNs. Aussi, les facteurs solubles CNTF (Ciliary Neurotrophic Factor, Facteur Ciliaire Neurotrophique) et LIF (Leukemia Inhibitory Factor, Facteur d'Inhibition de la Leucémie) induisent la voie STAT3/JAK pour permettre l'expression de GFAP (Glial Fibrillary Acidic Protein, Protéine Acide Fibrillaire Gliale)^{81,82}. Les récepteurs de CNTF et LIF permettent la phosphorylation de STAT3 en activant la kinase JAK (Janus Kinase), ce qui l'active et induit sa translocation au noyau pour activer l'expression de GFAP^{61,79,82}.

Les voies NOTCH, WNT, et BMP sont activées de façon concurrente pour stimuler la formation d'astrocytes. En effet, l'activation de la voie NOTCH dans les cellules non-neuronales permet l'expression de bHLH répresseurs qui empêchent l'expression de facteurs pro-neuronaux, ainsi que l'expression CSL-dépendante de GFAP^{80,83}. Effectivement, l'expression ectopique de HES5 promeut la gliogenèse aux dépens de la neurogenèse. De plus, la voie STAT3 induit la répression de facteurs de transcription bHLH neurogéniques en activant la voie NOTCH, tandis que la voie NOTCH active elle-même la voie STAT3 en recrutant la kinase JAK2^{61,82}. De son côté, BMP4 peut aussi activer la voie JAK/STAT pour promouvoir la maturation d'astrocytes GFAP⁺⁷⁷. Enfin, la voie WNT activée dans les neurones induit l'expression de ligands BMPs qui sont ensuite sécrétés et activent l'astrogenèse dans les cellules environnantes, tout en inhibant la différenciation oligodendrogliale^{61,77}.

1.3.2.d Oligodendrogenèse

Le dernier type cellulaire à être formé est l'oligodendrocyte. Alors que la voie NOTCH est réactivée pour former des astrocytes, il est nécessaire de l'inhiber de nouveau pour former les POCs car le maintien de l'activation de la voie NOTCH empêche la différenciation de POCs *in vitro* et la perte du signal NOTCH induit la maturation précoce de POCs^{61,83}. De plus, l'expression ectopique de SHH induit la formation de POCs exprimant des marqueurs spécifiques oligodendrocytaires comme OLIG1/2 *in vitro* et *in vivo*, tandis que sa perte entraîne une baisse de la population oligodendrocytaire et de la myélination⁸⁴⁻⁸⁷. Les deux facteurs de transcription OLIG1 et OLIG2 sont nécessaires et suffisants à la formation d'oligodendrocytes, et leur perte induit un défaut total de la production de POCs^{88,89}. Au contraire, l'activation de la voie BMP inhibe la formation de POCs *in vivo* et l'ajout de BMPs dans une culture de POCs induit même leur différenciation anormale en astrocytes GFAP⁺^{84,88,90,91}.

1.3.3 L'épigénétique dans le développement neural

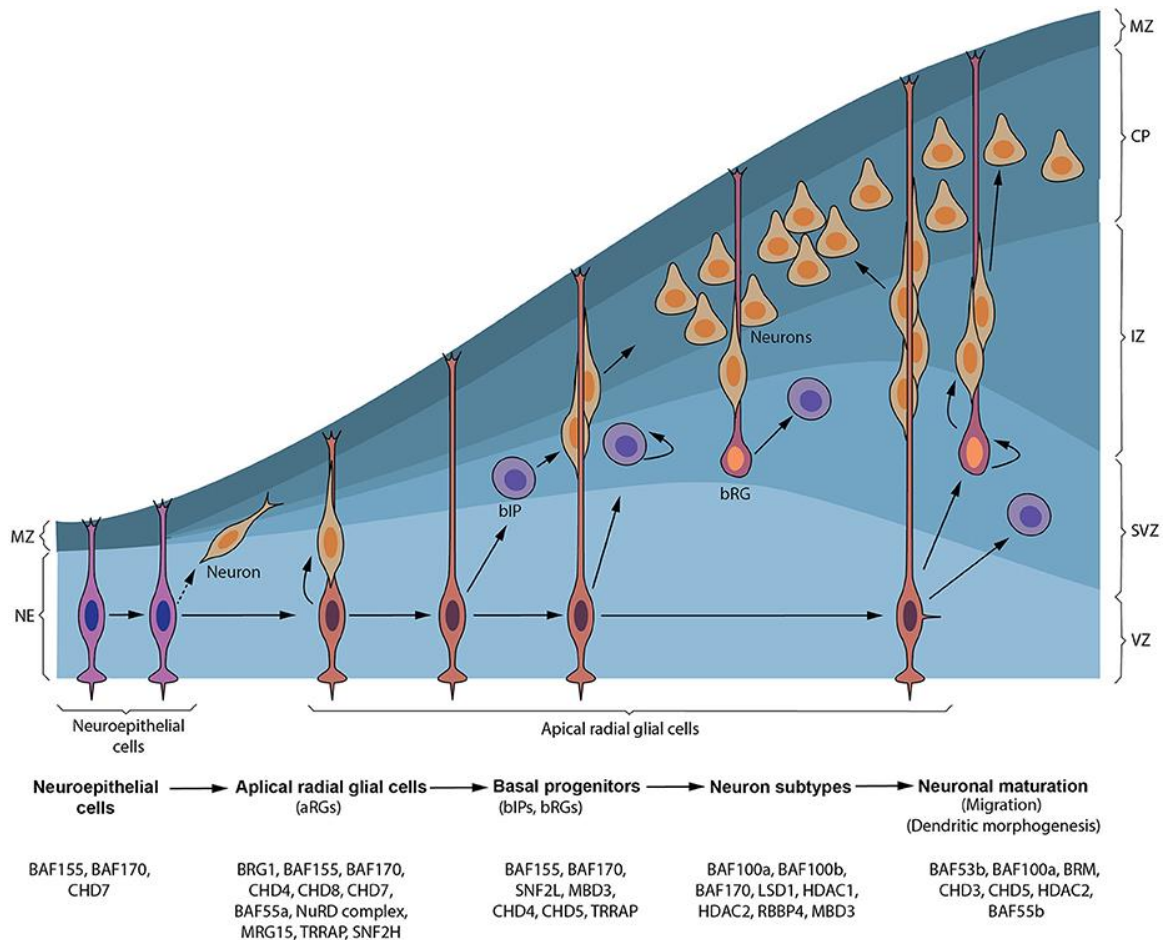


Figure 5: Les régulateurs de la chromatine sont impliqués dans le développement neural. (Sokpor et al, 2018). Des gènes impliqués dans la régulation de la structure de la chromatine ont été impliqués dans le développement du cortex cérébral. Chaque gène est associé à un phénotype pathogénique dans un modèle animal à un certain stade du développement.

Comme il a été établi ci-dessus, le développement neural nécessite l'expression de gènes selon un modèle spatio-temporel extrêmement précis. La modification du profil transcriptionnel cellulaire repose sur l'action de remodeleurs épigénétiques qui induisent des changements dans la chromatine pour activer ou inhiber les voies impliquées dans le neurodéveloppement. En effet, une multitude de facteurs impliqués dans les mécanismes épigénétiques énoncés précédemment ont une place pivotale dans l'identité des différents progéniteurs neuraux et leur maturation (Figure 5)^{92,93}. Particulièrement, ces facteurs agissent généralement en tant que complexes composés d'une multitude de protéines ayant des fonctions différentes, et dont la

composition change au cours du développement, leur octroyant une spécificité à la fois dans le ciblage et l'activité.

Le remodelage des nucléosomes de façon ATP-dépendante est composé de trois mécanismes: l'assemblage et l'organisation nucléosomale, l'accès à la chromatine, et l'édition nucléosomale. Les facteurs responsables de l'organisation nucléosomale permettent le positionnement correct des nucléosomes lors de la formation initiale de la chromatine. Les complexes capables de réaliser cet assemblage sont les familles ISWI (Imitation Switch) et CHD/NuRD (Chromodomain Helicase DNA-Binding / Nucleosome Remodeling Deacetylase)^{93,94}. Ils peuvent donc former des nucléosomes matures séparés de ~147pb. L'accès à la chromatine est ensuite modulé par la modification du positionnement de ces nucléosomes le long de la chromatine, soit en les déplaçant, soit en procédant au désassemblage de composantes du NPC ou du nucléosome en entier⁹³. La majorité des acteurs de l'accès à la chromatine est formée par la famille BAF (BRG1 or BRM Associated Factors) ou SWI/SNF (SWitch/Sucrose Non Fermentable). Généralement, l'agrégation de nucléosomes déplacés plus proches les uns des autres promeut l'éviction de dimères puis de l'octamère en entier, formant des régions d'ADN libres à la liaison par d'autres protéines^{93,95}. Enfin, l'édition de nucléosomes implique le changement de certaines histones pour des variantes, particulièrement par la famille INO80. La déstabilisation de nucléosomes par des CRC permet le retrait puis l'ajout subséquent de différentes histones, et donc la possibilité de rajouter de nouvelles histones à certains *loci*^{93,95}. Les complexes de remodelage de la chromatine étudiés ici sont le complexe BAF et le complexe NuRD.

1.3.3.a Le Complexe BAF

Le complexe BAF est composé de plus de 15 sous-unités, certaines étant mutuellement exclusives et occupant la même position à certaines étapes du développement, tandis que d'autres sont des composantes facultatives ou échangeables. Les différents complexes BAF sont nommés selon le moment où ils sont exprimés, soit pour le cortex le complexe esBAF (Embryonic Stem cell BAF), npBAF (Neural Progenitor BAF), et nBAF (Neuronal BAF). Il existe aussi le complexe PBAF (Polybromo-associated BAF) dont la composition est légèrement différente car ils ne contiennent que l'hélicase BRG1/SMARCA4 et possède les sous-unités

spécifiques PBRM1 (BAF180) et ARID2 (BAF200) contrairement aux complexes BAF qui comprennent à la place la sous-unité BAF250 (ARID1)^{93,93}. Le complexe BAF est composé d'une hélicase (SMARCA4 ou BRM/SMARCA2), d'un complexe-cœur formé de SMARCB1 (BAF47), SMARCC1 (BAF155), et/ou SMARCC2 (BAF170), et d'autres facteurs interchangeables⁹⁵. Chacune des sous-unités possèdent un ou plusieurs domaines capables de se lier à l'ADN directement ou à des histones, modifiées ou non, ou des domaines de liaison à d'autres protéines^{26,95}. Le complexe BAF a, sous toutes ses formes, une importance cruciale quant au développement neural, dans la formation de SCNs, leur prolifération, leur différenciation, et la maturation de types cellulaires ou de sous-types particuliers⁹⁵⁻⁹⁷.

La perte du complexe BAF dans un modèle animal ou cellulaire peut se faire soit par inactivation génétique de l'hélicase *SMARCA4* ou *SMARCA2*, ou par la double inactivation de *SMARCC1* et *SMARCC2* qui entraîne la séparation des sous-unités et leur dégradation subséquente⁹⁷. Ces modèles sont létaux au stade embryonnaire, et l'inactivation du complexe BAF SCN-spécifique induit la perte de structures corticales et la malformation du neuroépithélium olfactif⁹⁵⁻⁹⁷. Plus précisément, l'ablation du gène *Smarca4* après la formation des premières SCNs chez la souris entraîne la perte des progéniteurs Sox2+, notamment par la perte de la transcription Pax6-dépendante, ce dernier étant un marqueur exprimé majoritairement chez les PARGs et nécessaire à la formation du cortex et de l'oeil⁹⁷⁻¹⁰⁰. Dans le neuroépithélium olfactif, la perte de *Smarcc1* entraîne une baisse de la capacité proliférative des SCNs due à un arrêt du cycle cellulaire précoce, et la délétion de *Smarcc2* dans le cortex en développement induit un biais vers l'astrogenèse et une déplétion des progéniteurs radiaux^{96,98}. Outre *Smarca4*, l'inactivation des sous-unités du complexe npBAF *Phf10/Dpf2* (*Baf45a/d*) ou *Actl6a* (*Baf53a*) empêche aussi la prolifération et l'auto-renouvellement des SCNs, particulièrement en inhibant la voie Wnt dans la corticogenèse tardive¹⁰¹⁻¹⁰³. Enfin, la sous-unité spécifique au npBAF SS18 est nécessaire pour l'auto-renouvellement et sa délétion dans des SCNs est suffisante pour induire l'arrêt du cycle-cellulaire^{102,104}.

Pourtant, dans des progéniteurs plus précoces, le complexe BAF active l'expression de gènes gliaux comme *GLAST* ou *BLBP* et inhibe l'expression de gènes spécifiques de prolifération comme *Pax6* et *Sox2*, et le marqueur neuroépithélial *Ocln*, démontrant un pouvoir neurogénique^{101,105}. De plus, le complexe BAF est nécessaire pour l'inhibition de la voie Shh

dans le développement cortical en interagissant directement avec des facteurs Gli, et en se liant à des séquences cibles des facteurs Gli¹⁰⁶. Aussi, le complexe BAF est impliqué dans l'activation de la voie Shh et l'expression de ses gènes cibles dans des neurones en maturation^{103,106}. Additionnellement, Brg est impliqué dans l'expression des gènes de la voie Notch dans les SCNs, et donc dans l'inhibition de la différenciation neurale¹⁰¹. Ainsi, le complexe BAF exhibe une dichotomie dans ses fonctions cellulaires, étant à la fois promoteur de l'auto-renouvellement et inhibant la pluripotence.

Durant le développement, un changement critique du complexe npBAF au complexe nBAF est nécessaire pour la maturation neuronale. Ce transfert se fait par l'expression du répresseur *REST* dans les neurones post-mitotiques, qui réprime l'expression des micro ARNs mir-9* et mir-124¹⁰⁷. Par conséquent, la cible des ces miRNAs *Actl6a* est réprimé et la sous-unité exclusive *Actl6b* prend sa place dans le complexe nBAF^{107,108}. De plus, la maturation neuronale implique la compétition entre *Smarcc1* et *Smarcc2*, où le dernier est plus exprimé dans les progéniteurs puis son expression est diminuée lors de la différenciation^{109,110}. Dans le complexe BAF, ces protéines peuvent exister sous forme d'homodimère ou d'hétérodimère, donnant une spécificité au complexe par leur singularité structurale. En effet, SMARCC2 possède un domaine de liaison à REST à son extrémité N-terminale, contrairement à SMARCC1¹⁰⁹. L'augmentation graduelle de *Smarcc1* induit l'ouverture de la chromatine et l'activation de gènes cibles de Pax6, ainsi que le recrutement *Brg1*-dépendant des facteurs neurogéniques Sox11, Nfib et Pou3f4 aux cibles de Pax6 comme *Cux1* et *Tbr1*^{109,110}. Ensuite, *Actl6b* et le reste du nBAF (notamment les marqueurs de neurones matures *Ctip1/2* ou BAF100a/b, *Bcl11a/b*) sont nécessaires pour la maturation, la survie, et la dendrogenèse neuronale^{101,111}.

La maturation neuronale implique la migration neuronale vers les couches supérieures, un mécanisme affecté par la délétion de *Ctip1* à travers la perte de la répression du facteur *Sema3c*¹¹². De plus, *Ctip1* est nécessaire pour la morphogenèse de neurones de la moelle épinière¹¹³. Additionnellement, la dendritogenèse de neurones corticaux en culture est dépendante de l'expression d'*Actl6b*, notamment à travers un domaine qui lui est spécifique par rapport à son homologue *Actl6a*, avec qui il a pourtant 87% d'homologie¹⁰¹. Enfin, la transition vers le complexe nBAF implique l'échange de SS18 pour la sous-unité CREST (Calcium

Responsive Transactivator) à la sortie du cycle cellulaire¹¹². Cette dernière est responsable pour l'activation de *c-fos* lors de la synaptogenèse et est impliquée dans l'arborisation neuronale dépendante du calcium^{114,115}.

En plus de la neurogenèse, le complexe BAF est aussi impliqué dans la gliogenèse, particulièrement l'oligodendrogenèse. En effet, l'ablation de SMARCA4 ou SMARCC2 dans des SNCs adultes empêche la formation de neurones et induit directement la gliogenèse, et la perte de Smarca4 dans les SCNs murines induit la différenciation terminale des neurones avant la gliogenèse^{109,110,116}. De plus, lors de la maturation oligodendrocytaire, il y a un enrichissement significatif de l'ARN polymérase RNAPII au niveau des gènes de remodelage de la chromatine, particulièrement au *locus* de *Smarca4*¹¹⁷. La perte de Brg1 sous le contrôle du promoteur Olig1 induit des défauts de myélination bien que les OPCs soient toujours présents, indiquant un rôle du complexe BAF dans la maturation des oligodendrocytes^{117,118}. Enfin, le facteur de transcription Olig2 interagit avec Smarca4 pour recruter le complexe BAF au niveau des gènes responsables de la maturation oligodendrocytaire¹¹⁷. Pourtant, Smarca4 est positionné au promoteur du gène *Olig2* dans les progéniteurs neuraux, et induit sa répression¹¹⁶. Le complexe BAF est donc nécessaire pour la maintenance de la pluripotence à travers l'inhibition de la gliogenèse, mais d'autres facteurs induisant la différenciation en oligodendrocytes entraînent un changement d'environnement où le complexe BAF serait nécessaire pour leur maturation.

De façon intéressante, le complexe BAF interagit directement avec le complexe PRC1 (Polycomb Repressor Complex 1) et sa liaison entraîne l'éviction de PRC1 de la chromatine^{119,120}. Ainsi, un équilibre entre la répression par la marque épigénétique PRC1-dépendante H3K27Me3 et l'ouverture de la chromatine BAF-dépendante est indispensable pour le développement neural, l'activation de l'expression génique nécessitant à la fois le remodelage de la chromatine par le complexe BAF, et l'inhibition de la méthylation d'histone¹²⁰.

1.3.3.b Le Complexe NuRD

Le complexe NuRD est composé d'une entité centrale à fonction hélicase ATP-dépendante, traditionnellement CHD3 ou CHD4 (Mi-2 α/β) et dans certains cas CHD5, et d'une multitude d'autres sous-unités¹²¹⁻¹²⁴. *CHD4* est spécifiquement exprimé dans les progéniteurs neuraux précoces tandis que *CHD5* n'est seulement exprimé dans les précurseurs qui ont

commencé leur différenciation et *CHD3* dans les précurseurs terminaux dans les dernières étapes de la maturation neuronale¹²⁴. Les trois hélicases sont toutes présentes dans le cadre du complexe NuRD, mais elles sont mutuellement exclusives, et ont des fonctions singulières. En effet, l'ablation de *CHD4* sous l'action du promoteur Nestin induit une sortie prématurée du cycle cellulaire des SCNs et une réduction de la quantité de progéniteurs neuraux, et sa perte dans des cellules ES empêche la différenciation neuronale^{14,124,125}. En interagissant avec le facteur de transcription ZFX4, le complexe CHD4-NuRD peut réprimer l'expression de gènes neurogéniques¹²⁶. Au contraire, la fonction de *CHD5* se trouve dans la différenciation neuronale et la migration radiale initiale car sa perte entraîne un défaut de polarité dans les neurones sortant du cycle cellulaire^{124,127,128}. Enfin, *CHD3* est nécessaire pour la migration radiale plus tardive des neurones et n'est seulement exprimé dans les neurones post-mitotiques¹²⁴.

D'autres composantes du complexe NuRD sont des histones déacétylases: HDAC1 et HDAC2. Bien qu'elles possèdent 86% d'homologie au niveau de leur séquence protéique, leur expression est différente dans le développement neural. *Hdac1* est préférentiellement exprimé dans les progéniteurs neuraux, tandis que *Hdac2* est restreint aux neurones post-mitotiques^{129,130}. Pourtant, la perte de l'une des deux protéines ne cause pas de phénotype sévère, mais leur double inactivation empêche le développement cortical et est létal au stade embryonnaire¹²⁹. Notamment, la présence d'un seul allèle de *Hdac2* est nécessaire pour le développement neural normal, tandis qu'un allèle de *Hdac1* ne le permet pas, suggérant une redondance des fonctions de *Hdac1* par *Hdac2* avec une spécificité fonctionnelle additionnelle de *Hdac2*¹³¹. Cette singularité opère à travers la répression du gène codant pour la Protéine Kinase Delta *Pkc(δ)* par la liaison directe de Hdac2 au locus de *Prkcd*, sa perte induisant une augmentation locale activatrice de la marque H3K29Ac¹³¹. L'expression ectopique de *Pkcδ* entraîne une polarité aberrante et la déplétion de la population progénitrice neurale.

Les dernières sous-unités du complexe NuRD sont des protéines qui n'ont pas de fonction enzymatique. Au contraire, elles ont pour fonction de cibler le complexe à certaines régions du génome par leur capacité de se lier à l'ADN (MTA1/2/3), l'ADN méthylé sur la cytosine (MBD2 ou MBD3, MBD2 ayant une plus forte affinité pour l'ADN méthylé), ou l'histone H4 (Rbbp7/4)^{121,122}. La présence de ces sous-unités est nécessaire à la fois pour le fonctionnement du complexe, et pour son assemblage par leur fonction d'échafaud protéique

indispensable pour la formation du complexe¹²¹. Par exemple, l'invalidation génétique de *Mbd3* entraîne la dissociation du complexe NuRD et la perte de la répression des gènes *Nanog* et *Oct4*, laissant les cellules dans un état pluripotent perpétuel¹³². Contrairement à MBD2, MBD3 n'est pas enrichi au niveau des régions méthylées de l'ADN, mais est au contraire présent dans des régions portant des marques traditionnellement actives de la chromatine comme H3K27Ac¹³³. De plus, dans des cellules *Mbd3* déficientes, les régions normalement dites « bivalentes », sont déplétées de marques de répression H3K27me3 et enrichies à la place pour la marque H3K27Ac, démontrant un rôle pour le complexe de maintenance de l'état de préparation à la transcription¹³⁴. Dans les iPSCs (Cellules Souches Pluripotentes induites), la formation du complexe NuRD est dépendante de l'interaction avec le facteur de transcription Gatad2a, ou p66α, un facteur dont la déplétion dans des cellules différenciées n'a pas d'impact, mais qui est indispensable pour la répression de la pluripotence NuRD-dépendante¹³⁵.

1.3.3.c *TRRAP* et le complexe TIP60

Le gène *TRRAP* encode une protéine de 437kDa qui appartient à la famille des kinases PIKK (phosphoinositide 3-kinase-related kinase) qui n'a pas d'activité enzymatique, mais qui est extrêmement bien conservée jusqu'à la levure¹³⁶. Cependant, elle est présente en complexe avec une multitude de facteurs épigénétiques et de facteurs de transcription, et sa présence est indispensable pour le développement^{136,137}. En tant que protéine échafaud dans les complexes HAT, l'homologue de *TRRAP* chez la levure *Tra1* est nécessaire pour la formation des complexes NuA4 et SAGA et leur stabilité, leur recrutement aux promoteurs, et leur activité acétyl-transférase¹³⁸. Le complexe NuA4/TIP60 est responsable de l'acétylation des histones H4 (K5, 8, 12) et H2A (K5) ainsi que ses variantes H2AZ et H2AX, mais aussi de plus d'une centaine de protéines grâce à sa sous-unité TIP60 (*KAT5*) qui possède l'activité acétyltransférase¹³⁹. En général, il est impliqué dans la régulation du cycle cellulaire par l'acétylation de protéines comme AuroraB, mais aussi à travers son interaction avec les facteurs Myc, E2F, et Rb¹³⁹⁻¹⁴². Sa présence est nécessaire pour la sortie du cycle cellulaire grâce à la répression de gènes de la mitose¹³⁹. Il joue aussi un rôle important dans la réparation des bris double-brins de l'ADN. En effet, le complexe reconnaît les marques H3K9me3 et H4K20me1/2 déposées aux *foci* de dommages à l'ADN et y recrute des facteurs de réparation des bris double brins^{143,144}. TIP60 est donc responsable de l'activation directe d'ATM (Ataxia Telangiectasa

Mutant) par son acétylation, et l'acétylation de p53 lors de l'arrêt du cycle cellulaire est aussi dépendante de TIP60^{144,145}. En plus de son rôle d'acétyltransférase, TIP60 est en complexe avec l'hélicase ATP-dépendante p400 et la chaperonne de l'histone H2A.Z et est essentiel au fonctionnement de ce complexe en tant qu'échangeur de variantes d'histones¹⁴⁶⁻¹⁴⁹.

Dans les ESCs, le complexe Tip60 agit de façon indépendante de son activité acétyltransférase pour activer leur auto-renouvellement. Effectivement, son interaction ESC-spécifique avec Hdac6 semble lui donner un rôle répresseur où son recrutement au niveau des promoteurs de gènes de différenciation comme *Nodal*, *Dkk4*, ou *Gata6* induit la déacétylation de leurs *loci*¹⁵⁰. Ensuite, l'exclusion nucléaire de Hdac6 pourrait être le signal d'activation des gènes de différenciation, pour lesquels la transcription nécessite l'activité acétyltransférase de Tip60¹⁴⁶. De plus, la déposition de l'histone H2A.Z aux *loci* des gènes cibles de la voie NOTCH est TIP60-p400 dépendante, et celle-ci diminue leur expression^{147,151}. L'association de TIP60 avec p400 inhibe son activité HAT, tandis que sa dissociation lui permet d'acétyler les variantes H2A.Z et d'activer la voie moléculaire NOTCH^{147,148}. Chez la drosophile, l'ablation du complexe Tip60 dans les SNCs entraîne une perte de la polarité de façon aPKC dépendante et des anomalies du fuseau mitotique, et par conséquent la déplétion des progéniteurs neuraux due à une réduction des divisions asymétriques¹⁵². Additionnellement, Pax6 interagit directement avec Tip60 dans les progéniteurs neuraux de la rétine, et cette liaison est essentielle pour son activité transcriptionnelle¹⁵³. De plus, l'homologue chez la drosophile de BAF53a, Bap55, est retrouvé à la fois dans le complexe BRM et le complexe Tip60, où il est nécessaire pour la dendrogenèse des neurones olfactifs¹⁵⁴. Enfin, la fonction HAT de Tip60 est indispensable pour la neuroplasticité et la mémoire chez la drosophile¹⁵⁵⁻¹⁵⁷.

La protéine TRRAP fait aussi partie des complexes STAGA contenant les HATs GCN5 ou PCAF, et importants pour le développement neural. En effet, l'invalidation génétique de *Gcn5* induit l'exencéphalie dans un modèle murin et empêche la fermeture du tube neural, par l'inhibition de la prolifération des progéniteurs neuraux¹⁵⁸⁻¹⁶⁰. Plus précisément, STAGA interagit avec N-Myc à travers des sites de liaison de TRRAP et GNC5, lui permettant d'être localisé aux cibles de Myc dans les CSNs et d'induire leur acétylation et leur expression¹⁶⁰⁻¹⁶². Les gènes affectés par la perte de N-Myc ou GCN5 sont impliqués dans la prolifération des CSNs, la gliogenèse, et la guidance axonale¹⁵⁹. Dans la rétine, le complexe STAGA interagit

directement avec CRX (Cone Rod Homeobox) pour acétyler les *loci* de ses cibles et maintenir leur expression, mécanisme essentiel pour la survie de ces cellules¹⁶³. La perte de *Pcaf* n'est pas délétère, mais l'ablation de *Gcn5* en conjonction de celle de *Pcaf* donne lieu à un phénotype plus sévère que celle de *Gcn5* seul, suggérant un rôle redondant pour *Pcaf*¹⁶⁴. En revanche, *Pcaf* est nécessaire pour la régénération axonale dans le SNC, grâce à son activité acétyltransférase déposant la marque H3K9Ac dépendante de l'activation de la voie ERK¹⁶⁵. *Gcn5* et *Trrap* sont notamment indispensables pour la transcription des gènes cibles de la famille de facteurs de transcription E2F¹⁶⁶.

TRRAP est une protéine échafaud dont la présence est essentielle pour la formation d'une multitude de complexes à activité acétyltransférase. Particulièrement, TRRAP est recruté à la chromatine en premier, puis les complexes sont formés dessus, démontrant un recrutement TRRAP-dépendant aux *loci* ciblés afin de les acétyler¹⁶⁷. Le gène *Trrap* a été impliqué dans la régulation de multiples voies moléculaires, et est particulièrement essentiel à l'activation de la voie Notch dans le développement de l'aile de la *Drosophile*¹⁶⁸. De plus, TRRAP fait partie d'un complexe capable de dimétyler et trimétyler H3K79 en conjonction avec β -Caténine de façon à réguler négativement les cibles en aval de la voie WNT¹⁶⁹. La régulation de la β -Caténine se fait aussi directement par TRRAP qui induit son délogement de la chromatine ainsi que son ubiquitination et sa dégradation subséquente¹⁷⁰. Enfin, TRRAP est impliqué dans la régulation de la voie Ras chez le ver *C. Elegans*¹⁷¹. La délétion du gène *Trrap* est létal au stade embryonnaire, et son ablation sous le promoteur Nestin cause des défauts du cycle cellulaire, une perte de la polarité cellulaire, et une différenciation prématurée des progéniteurs neuraux^{130,172,173}. TRRAP et GCN5 sont aussi nécessaires pour l'activation de la transcription par le complexe TFTC (TBP Free TAF Complex) qui recrute POLRII pour initier la transcription¹⁷⁴.

D'autres histone acétyltransférases sont aussi impliquées dans le développement neural. Notamment, les HATs P300 et CBP sont importantes pour l'acétylation de l'histone H3. Particulièrement, des études dans le modèle *Xenopus laevis* ont démontré que l'ablation de *p300/CBP* entraîne la neuralisation de l'embryon en entier; toutes les cellules souches embryonnaires ont une propension intrinsèque à la neurogenèse, et l'inhibition de ce mécanisme est acétylation-dépendante¹⁷⁵. Des mutations hétérozygotes de *Cbp* ne sont pas suffisantes pour

affecter le développement morphologique murin, mais entraînent des défauts dans le comportement, et les embryons hétérozygotes pour *p300* ne présentent un phénotype exencéphalique restreint et avec une moindre pénétrance en comparaison avec les invalidations homozygotes de chacun des gènes^{176,177}. Cependant, la double perte hétérozygote des deux gènes est létale au stade embryonnaire, suggérant une redondance fonctionnelle des deux HATs, et un effet dose-dépendant de l'acétylation sur le développement. L'ablation hétérozygote de *Cbp* dans des progéniteurs neuraux induit la perte de leur capacité à se différencier en neurones, astrocytes, et oligodendrocytes¹⁷⁷. Notamment, Cbp permet l'acétylation H3K9/14 au niveau des promoteurs de *$\alpha 1$ -tubulin*, *Gfap*, et *MBP* (Myelin Binding Protein), des marqueurs neuronal, astrocytaire, et oligodendrocytaire respectivement. Cbp interagit avec Stat3 au niveau du promoteur de *Gfap* pour induire l'astrogenèse¹⁷⁸. Ensuite, l'expression de *Hdac3* est suffisante pour déplacer Stat3 et le nouveau complexe Hdac3/Cbp est nécessaire pour l'inhibition de l'expression de gènes astrogéniques (*Gfap*, *NFIA*, *Aqp9*) et la nouvelle localisation de Cbp au locus de gènes de l'oligodendrogenèse comme *Olig2* et *Smarca2*¹⁷⁹. Le facteur de transcription Ngn1 (homologue murin de NEUROG1) active la neurogenèse notamment en séquestrant le complexe Cbp/Smad1 hors des promoteurs de gènes astrocytaires et en inhibant les facteurs de transcription Stat¹⁸⁰.

1.4. Retard du développement, déficit intellectuel, et autisme

1.4.1 Définitions et informations générales

Les maladies neurodéveloppementales (Neurodevelopmental Disabilities – NDDs) sont relativement prévalentes dans la population; on estime qu'entre 3 et 5% de la population est affectée d'une forme de NDD^{181–183}. Elles peuvent être séparées en trois diagnostics: Retard du Développement Global (RDG ou DD, délai du développement), Déficit Intellectuel (DI), et Trouble du Spectre de l'Autisme (TSA). Le RDG est défini par un déficit dans au moins deux domaines du développement, soit la motricité globale ou fine, le langage, les capacités cognitives ou sociales, ou les tâches de quotidien¹⁸³. Il se présente généralement avant l'âge de 5 ans et est le présage d'un DI. Le DI ne peut, par définition, être diagnostiqué qu'à partir de 5 ans et est traditionnellement caractérisé grâce à un test de QI afin d'analyser les capacités de l'enfant dans l'intelligence et le comportement adaptatif, dans un contexte social déterminé au

préalable. On diagnostique un DI quand le score QI de l'individu est en dessous de 70¹⁸¹. Le TSA est plus compliqué à diagnostiquer et à distinguer des autres NDDs, du fait de ses symptômes parfois plus légers ou confondus avec des domaines impliqués dans le RGD ou le DI. Un enfant présentant avec un TSA a des défauts dans les relations interpersonnelles, la communication sociale, et parfois des comportements répétitifs ou des intérêts restreints. De plus, 70% des individus ayant un TSA ont une forme de DI, et 10% des cas de DI sont associés à un TSA¹⁸².

En général, la génétique des NDDs est mal comprise, bien qu'un lien causal entre anomalies chromosomiques, variabilité du nombre de copies, ou mutations ponctuelles ont été établis pour une multitude de phénotypes. Les mâles ont une prévalence de TSA au moins trois fois plus haute que les femelles, potentiellement expliqué par un modèle d'effet protecteur féminin où des variants ayant une charge pathogénique plus forte sont nécessaires pour un diagnostic, dû à des causes à la fois sociales et génétiques du fait de la présence d'un nombre de gènes associés au TSA sur le chromosome X^{184,185}. Aujourd'hui, un diagnostic ne se fait plus simplement par le phénotype, mais aussi par le génotype, où les individus sont aussi caractérisés par leur gène affecté grâce au séquençage de l'exome ou du génome¹⁸³. Il y a un enrichissement significatif de variants tronquant la protéine (VTPs) dans les cohortes de TSA/DI, et ce seulement dans les gènes qui ont une probabilité d'intolérance à la perte de fonction (pLI – probability of loss-of-function intolerance) qui est haute (pLI > 0.995), c'est-à-dire les gènes sensibles à l'haploinsuffisance¹⁸⁵. Elle peut avoir lieu par la perte de fonction d'un domaine protéique, ou par la perte de l'expression du gène par la dégradation des ARNm non-sens (DMN). Pourtant, la majorité des variants identifiés (~65 % dans les trios *de novo*, les trios avec héritage, et les cas-contrôles *de novo*) sont des mutations faux-sens ponctuelles, laissant penser que ceux-ci fonctionnent par un mécanisme autre que l'haploinsuffisance où des domaines particuliers de la protéine sont affectés^{185,186}. Malheureusement, des mutations faux-sens peuvent être localisées dans des régions du gène n'ayant pas d'impact, rajoutant du bruit dans les contrôles auxquels les individus affectés, et des mutations rares peuvent être détectées chez l'individu sans qu'elles ne soient causatives du phénotyp. Aussi, des mutations dans différentes régions du gène peuvent causer des phénotypes à pénétrance variable à cause de leur impact unique sur la fonction de la protéine. Ainsi, les méthodes utilisées aujourd'hui ne permettent

qu'à la caractérisation de mutations causant des phénotypes très sévères, et des cas peuvent être ajoutés à une cohorte quand un phénotype plus léger est diagnostiqué avec une mutation dans le même gène.

Les gènes dont l'association avec un NDD est significative peuvent être regroupés dans des groupes fonctionnels distincts, soit la fonction neuronale/synaptique, le cytosquelette et les molécules d'adhésion, la régulation de l'expression génique, et la régulation de la traduction ou dégradation protéique^{182,185}. Dans une étude de 35,584 individus diagnostiqués avec un TSA avec ou sans co-morbidités (DI/DD), 99 gènes ont été identifiés comme étant statistiquement sur-représentés dans les variants identifiés chez les individus affectés. Sur les 99, 98 sont enrichis dans le cortex et ils sont tous exprimés en plus grande quantité dans le cortex prénatal par rapport au cortex postnatal¹⁸⁵. De plus, les types cellulaires associés à ces gènes sont les neurones en maturation et matures, les astrocytes, et les POCs. En séparant les gènes impliqués dans la fonction neuronale et ceux impliqués dans la régulation de l'expression génique, cette étude démontre que les gènes de régulation de l'expression génique sont plus exprimés dans les stades primaires du développement, et sont enrichis dans le TSA associé au DI/DD; les gènes de la fonction neuronale sont au contraire enrichis plus tard dans le développement et impliqués dans le TSA seul.

1.4.2 Caractérisation des phénotypes étudiés

1.4.2.a Les BAFopathies et *SMARCC2*

Les syndromes causés par des mutations dans les gènes codant pour des sous-unités du complexe BAF, et plus récemment aussi dans le gène codant pour le facteur de transcription SOX11, ont été englobés dans un spectre de pathogénies nommé Troubles de Déficience Intellectuelle liés au SWI/SNF (TDISS)¹⁸⁷. Les TDISS représentent un spectre continu de phénotypes d'abord appelés Syndrome de Coffin-Siris (CSS, [OMIM: PS135900]) et Syndrome de Nicolaides-Baraitser (NBS, [OMIM: 601358]), mais dont la pénétrance extrêmement variable associée à un même gène a permis un regroupement sous une même identité plutôt que des syndromes singuliers. Les caractéristiques principales du SCS sont le DI, l'hypotonie, des problèmes d'alimentation, l'hypertrichose, et la malformation d'au moins un ongle et/ou un doigt de la main¹⁸⁸. Le SNB comprend le DI, une petite taille, la brachydactylie, des problèmes

du comportement, et des crises épileptiques¹⁸⁹. Nous avons énuméré les caractéristiques des phénotypes associés à chaque gène, et montré qu'il existe une grande variabilité dans les présentations cliniques des individus affectés¹⁹¹.

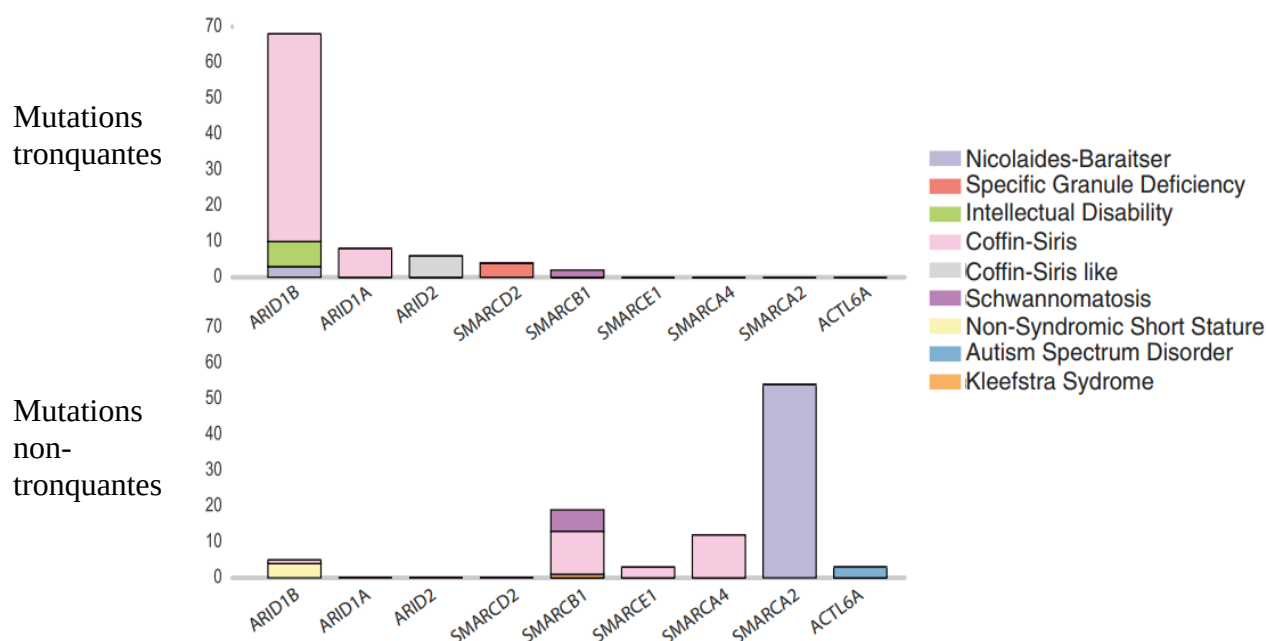


Figure 6: Répartition des mutations associées à un syndrome neurodéveloppemental dans les gènes du complexe BAF (Adapté de Mashtalir, et al. 2018). Des mutations tronquantes (mutations non-sens ou dans un site d'épissage) ou non-tronquantes (mutations ponctuelles faux sens ou additions/délétions ne changeant pas le cadre de lecture) ont été associées aux TDISS et d'autres maladies). Les barres représentent le nombre d'individus rapportés dans la littérature pour chacun des syndrômes.

Dans tous les cas reportés de TDISS, deux tiers des mutations sont retrouvés dans les gènes *ARID1B* (*BAF250b*) et *SMARCA2* (*BRM*, *BAF190B*) (Figure 6)¹⁸⁷. En revanche, 97% des mutations probablement pathogéniques dans le gène *ARID1B* (n=119) fonctionnent par une perte de fonction (changement du cadre de lecture, troncation, site d'épissage). Au contraire, les mutations de *SMARCA2* sont presque toutes faux-sens ou ne changent pas le cadre de lecture, avec une accumulation significative au niveau du domaine SNF2 ATPase, pointant vers un mécanisme pathogénique différent. Cependant, il est impossible de savoir si la cause de la prévalence des mutations de *ARID1B* est une mutagénicité particulière à son locus, ou une pathogénicité mutationnelle moins restrictive. La famille des gènes SMARC semble être plus affectée par des mutations ponctuelles agissant par un mécanisme de gain de fonction où ces

mutations sont localisées dans des domaines fonctionnels de la protéine. Les gènes de la famille ARID sont plutôt porteurs de mutations qui ne sont pas regroupées et qui fonctionneraient par perte de fonction ou haploinsuffisance.

Nous avons présenté 15 individus portant des mutations *de novo* dans le gène *SMARCC2* et présentant des caractéristiques du TDISS¹⁹⁰. Les individus ont tous une forme de DI allant de léger (5/15) à modéré/profond (10/15), avec 13 des individus ayant un défaut du langage. Huit des individus sont porteurs d'une mutation faux-sens et un individu porte une délétion d'un codon, deux individus sont porteurs de mutations tronquantes dans le N-terminus de la protéine, et quatre individus sont porteurs de mutations dans des sites d'épissage. Toutes les mutations sont localisées au niveau de nucléotides extrêmement conservés dans les vertébrés, et semblent indiquer un rôle important des domaines protéiques SWIRM et SANT de *SMARCC2* dans la pathologie.

1.4.2.b Syndrome Neurodéveloppemental associé à *CHD3*

Plusieurs des hélicases de la famille des CHD ont été associées à des syndromes neurodéveloppementaux. Ces syndromes impliquent tous une forme de DI ou de TSA, et la majorité implique des anomalies du SNC (macro/microcéphalie, anomalies du prosencéphale). Les phénotypes variables peuvent être expliqués par les rôles à la fois redondants et singuliers des hélicases dans le développement¹⁹¹.

Tableau I - Syndromes associés aux hélicases de la famille CHD

| Gène | Phénotype | Références |
|------|--|------------|
| CHD1 | Syndrome de Pilarowski-Bjornsson [OMIM:617682] | 192 |
| CHD2 | Épilepsie myoclonono-astatique du jeune enfant [OMIM:615369] | 193–196 |
| CHD3 | Syndrome de Snijders Blok – Campeau [OMIM:618205] | 197,198 |
| CHD4 | Syndrome de Sifrim-Hitz-Weiss [OMIM:617159] | 199,200 |
| CHD5 | - | |
| CHD6 | - | |
| CHD7 | Syndrome CHARGE [OMIM:214800] | 201–203 |
| CHD8 | Autisme [OMIM: 615032] | 204,205 |
| CHD9 | - | |

Nous avons rapporté 35 individus avec des mutations *de novo* dans le gène *CHD3*¹⁹⁷. Tous les individus ont un DI avec un QI allant de 70-85 à en dessous de 35, avec des défauts dans l'apprentissage du langage. La majorité des individus affectés présentent une macrocéphalie, tandis qu'un présente une microcéphalie; 75% ont aussi de l'hypotonie. Sur les 35 individus, 31 sont porteurs de mutations faux-sens, deux sont porteurs d'une délétion d'un codon et un porteur d'une mutation d'épissage ne changeant pas le cadre de lecture. Le dernier individu est porteur d'une mutation à l'extrémité C-terminale changeant le cadre de lecture et rajoutant 108 acide-aminés. Ce transcrit est correctement exprimé et n'est pas dégradé par DMN.

1.4.2.c Syndrome Neurodéveloppemental associé à *TRRAP*

Les KAT/KDACs sont impliqués dans une multitude de syndromes neurodéveloppementaux et neurodégénératifs. Les phénotypes associés sont multiples et à sévérité variable.

Tableau II - Syndromes Associés aux KATs/KDACs

| Gène | Phénotype | Références |
|--------------|---|------------|
| HDAC4 | Délétion Chromosome 2q37 [OMIM:600430] | 206 |
| HDAC6 | Chondrodysplasie avec Platyspondylie, Brachydactylie, Hydrocéphalie, et Microphthalmie [OMIM:300863] | 207,208 |
| HDAC8 | Syndrome Cornelia De Lange [OMIM:300882] | 209,210 |
| HDAC9 | Autisme, Schizophrénie | 211–213 |
| KAT6A | Retard du Développement global et microcéphalie [OMIM:616268] | 214,215 |
| KAT6B | Syndrome Génitopatellaire [OMIM:606170] Syndrome de Say-Barber-Biesecker-Young-Simpson [OMIM:603736] | 216–219 |
| CREBBP (CBP) | Syndrome de Rubinstein-Taybi 1 [OMIM:180849] Duplication Chromosome 16p.13.3 [OMIM:613458] | 220–222 |
| EP300 (p300) | Syndrome de Rubinstein-Taybi 2 [OMIM:613684] | 223,224 |
| BRPF1 | DI avec Facies Dysmorphique et Ptosis [OMIM:617333] | 225 |

Nous avons caractérisé 24 individus portant des mutations faux-sens *de novo* dans le gène *TRRAP*²²⁶. Les individus présentent tous un DI ou une forme de TSA, mais il semblerait y avoir une corrélation entre la localisation de la mutation et la sévérité du phénotype, avec une accumulation significative de mutations autour de la mutation récurrente NM_001244580.1 c.3127G>A, p.A1043T, donnant lieu à un phénotype syndromique associé à des malformations du cerveau, du coeur, des reins, ou du système génito-urinaire. Le second phénotype est caractérisé par un DI ou TSA avec ou sans épilepsie.

1.5. Gènes à l'étude

1.5.1 SMARCC2: SWI/SNF Related, Matrix Associated, Actin Dependent Regulator of Chromatin Subfamily C Member 2

Le gène *SMARCC2* encode la protéine BAF170, une protéine faisant partie du cœur du complexe BAF. Récemment, une étude de la composition structurale du complexe BAF a permis

la caractérisation systématique des interactions entre ses sous-unités²²⁷. Par la fixation et la caractérisation des différents complexes BAF, Mashtalir et al. ont pu démontrer que le cœur initial est formé d'un dimère SMARCC (homo/hétérodimère avec SMARCC1 et/ou SMARCC2) et d'une sous-unité SMARCD.

L'intégration d'autres sous-unités structurales se fait de façon subséquente, tandis que le module ATPase formé au minimum de l'hélicase SMARCA2/4, d'une sous-unité ACTL6, d'actine, et de BRD7 ne se fait qu'en dernier.

Les mutations que nous avons identifiées dans le gène *SMARCC2* sont localisées majoritairement dans des domaines connus, soit les domaines SMARCC et le domaine SANT. Les mutations sont visualisables dans la section résultats 3.2.2 (Figure 13). Les premiers, SMARCC_N localisé en N-terminal, et SMARCC_C, en C-terminal, sont des domaines présents dans BAF155/170 et qui sont extrêmement conservés jusqu'à leurs orthologues chez *S. Pombe* Ssr1/2 et chez *S. Cerevisiae* Swi3, mais dont la fonction n'est pas connue^{228,229}. Le second domaine touché est le domaine SANT (Swi3, Ada2, N-CoR, TFIIB), qui est un court domaine retrouvé chez plusieurs protéines impliquées dans le remodelage de la chromatine ou la régulation de l'expression génique. Il a une homologie structurale avec les domaines HTH (helix turn helix) qui se lie à l'ADN, mais a au contraire une fonction de liaison aux protéines, notamment aux histones^{230,231}. Un dernier domaine est présent dans la protéine SMARCC2: le domaine SWIRM. Celui-ci est un domaine associé à la liaison à l'ADN et aux nucléosomes, et est indispensable pour la formation des complexes associés aux protéines ayant ce domaine^{232,233}.

SMARCC1 et SMARCC2 sont interchangeable dans le complexe BAF, étant capables de former des homo/hétéro-dimères, mais leur stœchiométrie change au cours du développement, suggérant à la fois une redondance et un rôle unique pour chacune des protéines. Par exemple, le complexe esBAF présent dans les cellules souches embryonnaires ne contient que l'homodimère SMARCC1²³⁴. La double délétion de *SMARCC1/SMARCC2* induit la dégradation quasi-totale des autres sous-unités du complexe BAF, ce qui signifierait que la formation du complexe est dépendante du dimère SMARCC et que les autres sous-unités sont dégradées de façon ubiquitaire si elles ne sont pas incorporées dans un complexe²²⁷.

1.5.2 CHD3: Chromodomain Helicase DNA binding protein 3

Le gène *CHD3* encode une hélicase ATP-dépendante, qui avec CHD4 et CHD5, peut s'insérer dans le complexe NuRD et agir en tant que remodelleur de la chromatine. Nous avons

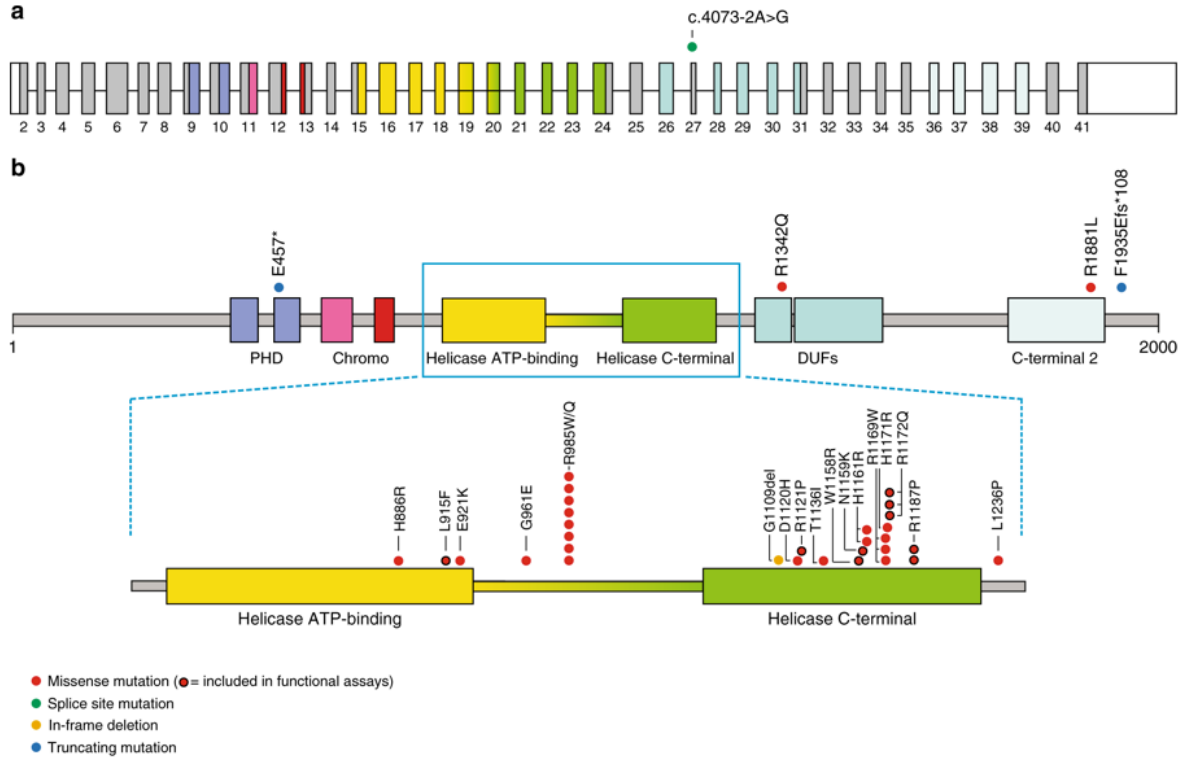


Figure 7: Localisation des mutations dans *CHD3* associées au syndrome Snijders-Blok-Campeau. Chaque individu est représenté par un point, et chaque mutation est classifiée selon sa fonction : le rouge représente une mutation faux-sens, le vert une mutation d'un site d'épissage, le jaune une délétion n'affectant pas le cadre de lecture, et le bleu une mutation tronquante.

caractérisé 35 individus avec 23 mutations *de novo* différentes; toutes les mutations sauf quatre étant des mutations faux-sens¹⁹⁷. La localisation des mutations présente un regroupement frappant, avec la quasi-totalité des mutations retrouvés dans les domaines de l'hélicase de liaison à l'ATP ou C-terminal, ou entre les deux (Figure 7). Ce phénomène suggère que la pathogénicité des mutations est due à une perte de la capacité de remodelage de la chromatine, soit par sa capacité de lier et d'hydrolyser l'ATP, soit par sa capacité de remodelage des nucléosomes.

La protéine CHD3 possède plusieurs domaines connus, dont deux domaines PHD (Plant HomeoDomain) qui sont impliqués dans l'épigénétique et la régulation de la transcription. Dans la HAT p300, le domaine PHD est responsable de sa liaison aux nucléosomes hautement

acétylés en conjonction avec son domaine Bromo²³⁵. Le domaine Chromo (Chromatin Organization Modifier) participe à la formation de l'hétérochromatine et pourrait prendre part dans la reconnaissance de nucléosomes méthylés²³⁶. Il est intéressant de noter que CHD3 et CHD4 sont particulièrement présents au niveau de l'hétérochromatine, et CHD3 est significativement enrichi dans l'hétérochromatine périnucléolaire¹²⁵. Tous ces domaines sont aussi présents dans son homologue CHD4, où les domaines PHD, Chromodomain, DUF (Domain of Unknown Function), et domaine C-terminal sont nécessaires pour réguler leur activité ATPase^{132,237}.

1.5.3 TRRAP: Transformation/Transcription domain Associated Protein

Le gène *TRRAP* encode une protéine de 3859 acides aminés de la famille des PIKKs qui n'a pas de fonction enzymatique à cause de la perte d'un acide aminé critique dans son site catalytique²³⁸. Toutes les PIKKs ont un domaine kinase, flanqué de domaines FAT et FATC, de sorte à former un motif FAT/PIKK/FATC. Les domaines FATC des PIKKs sont nécessaires pour l'interaction avec TIP60 et sont fonctionnellement interchangeables²³⁹. Le domaine FAT, quant à lui, est relativement conservé mais sa fonction n'est pas connue²⁴⁰.

Nous avons caractérisé 24 individus portant 17 mutations *de novo* différentes dans le gène *TRRAP*²²⁶. Pourtant, toutes sauf une mutation sont localisées en dehors des domaines connus, et sont restreints de façon significative dans des régions spécifiques, suggérant la présence de domaines fonctionnels non-identifiés (Section résultats 3.1.2, Figure 8). Effectivement, dans son homologue chez la levure *Tra1*, la délétion systématique des séquences répétées de type HEAT (Huntingtin, EF3, PP2A, TOR1 – structure secondaire formée de deux hélices alpha reliées par une liaison courte) a démontré que les régions mutées chez les individus les plus affectés sont indispensables pour la survie cellulaire et la liaison avec les complexes SAGA et/ou NuA4¹³⁸. En revanche, les délétions comprenant les mutations associées avec un phénotype moins sévère (acides aminés 1859-1932) sont viables et permettent quand même l'association avec les complexes SAGA et NuA4, suggérant un mécanisme pathogénique différent.

1.6 Problématique et Hypothèses

Problématique : Les régulateurs de la chromatine sont importants pour le développement neural par leur coordination des changements de profils d'expression génique au cours du temps à travers des modifications directes de la chromatine qui lui donnent une structure plus ou moins ouverte et prédisposée à la transcription génique. Nous avons caractérisé des individus présentant des symptômes de syndrômes neurodéveloppementaux caractérisés principalement par un DI, DD, et/ou TSA. Ceux-ci sont porteurs de variants dans des gènes codant pour des régulateurs de la chromatine (*TRRAP*, *SMARCC2*, et *CHD3*).

Hypothèse : Nous émettons l'hypothèse que les mutations ont un impact sur la fonction de ces sous-unités de complexes épigénétiques, qui se reflète par des changements de l'expression de gènes normalement régulée par ces complexes, et que c'est ce dérèglement qui donne lieu aux syndrômes observés.

Objectifs : Afin de tester l'impact des variants potentiellement pathogéniques sur le fonctionnement des complexes associés à *TRRAP*, *SMARCC2*, et *CHD3*, on cherche à évaluer si l'expression génique globale de cellules provenant d'individus porteurs de ces variants est significativement différente de celles provenant d'individus sains. Pour tester cette hypothèse, nous avons utilisé des cellules provenant de certains des individus affectés, soit des lignées primaires fibroblastiques et des lignées transformées lymphoblastoïques afin d'étudier l'impact des mutations sur l'expression génique par séquençage à haut débit de l'ARNm. Ensuite, sachant que la régulation de l'expression génique se fait par des modifications de la chromatine orchestrées par les complexes de remodelage de la chromatine que l'on étudie, et que celles-ci ont un impact sur l'ouverture de la chromatine, nous cherchons à évaluer cet état par ATACseq (Assay of Transposase Accessible Chromatin).

Résultats attendus : Nous nous attendons à trouver des profils d'expression génique différents dans les cellules provenant d'individus affectés par rapport aux cellules d'individus sains, et que les gènes affectés par des variants dans un même gène soient en majorité identiques car ce seraient les gènes cibles du complexe en temps normal. De plus, nous nous attendons à ce que l'ouverture de la chromatine soit différente entre les individus affectés et sains dans les régions proches des gènes dérégulés.

2. Matériel et Méthodes

2.1. Génération de lignées fibroblastiques et lymphoblastoïques et extraction d'ARN

2.1.1 Génération de lignées et extraction d'ARN

Les lignées fibroblastes utilisées sont établies depuis une biopsie de peau des individus. La biopsie est découpée et placée dans un pétri, peau vers le haut, dans un milieu DMEM (Dulbecco's Minimum Essential Medium) à haut glucose, 20% FBS (en anglais, Sérum Bovin Fétal), 1% Pénicilline/Streptomycine, 1% Glutamax. Huit jours plus tard et quand la biopsie a bien adhéré, les cellules sont trypsinisées puis du milieu frais est rajouté. Quand suffisamment de fibroblastes sont visibles autour de la biopsie, les cellules sont trypsinisées et transférées dans un nouveau contenant dans du DMEM à haut glucose, 10% FBS, 1% Pen/Strep, 1% Glutamax. Les cellules sont cultivées jusqu'à 80-90% confluence, puis trypsinisées, lavées deux fois au PBS 1X (en anglais, Saline Tamponnée au Phosphate), et lysées du Qiazol. L'ARN total est ensuite extrait à l'aide du kit Qiagen RNEasy mini, et traité à la DNase avec le kit Turbo DNA free.

Les lignées lymphoblastoïques utilisées sont établies à partir du sang des individus. Six à huit mL sang sont déposés sur un volume 1:1 équivalent de Ficoll-Plaqué dans un tube de 15mL qui est centrifugé sans freinage à 2000rpm pendant 20 minutes. Les cellules mononucléées sanguine périphériques (Peripheral Blood Mononucleated Cells, PBMCs) à l'interface entre le plasma et les cellules erythroïdes sont transférées dans un nouveau tube et lavées plusieurs fois avec du milieu RPMI (Roswell Park Memorial Institute) avec 1% FBS. Les cellules sont ensuite resuspendues dans du RPMI contenant du virus Eppstein-Barr et incubées à 37°C pendant une heure. Enfin, elles sont transférées dans un contenant T75 avec 7mL de RPMI 10% FBS, 1% Pen/Strep, 1% Glutamax, 0.5gml⁻¹ Cyclosporine A. Quand les cellules ont atteint une confluence optimale (un million de cellules par mL de milieu), les cellules sont centrifugées et lavées deux fois au PBS. L'ARN total est isolé de la même façon que pour les fibroblastes.

Tableau III - Cellules utilisées dans les expériences.

| Gène Affecté | Variant | Type Cellulaire | Sexe | Age |
|---------------------------------|---------------------------|----------------------------------|------|--------|
| CHD3 | NM_001005273.2 : p.L915F | Fibroblaste | F | 4 ans |
| CHD3 | NM_001005273.2 : p.N1159K | Fibroblaste | F | 8 ans |
| SMARCC2 | NM_003075.3 : p.L610P | Fibroblaste | M | 11 ans |
| SMARCC2 | NM_003075.3 : p.N134D | Fibroblaste | M | 5 ans |
| SMARCC2 | NM_003075.3 : c.1833+1G>T | LCL (pas de RNAseq) | M | 11 ans |
| TRRAP | NM_001244580.1 : p.A1043T | Fibroblaste (pas de RNAseq), LCL | F | 2 ans |
| TRRAP | NM_001244580.1 : p.W1866C | Fibroblaste | F | 10 ans |
| TRRAP | NM_001244580.1 : p.L805F | Fibroblaste, LCL | F | 8 ans |
| TRRAP | NM_001244580.1 : p.G1883R | Fibroblaste (pas de RNAseq) | M | 16 ans |
| TRRAP | NM_001244580.1 : p.A1043T | LCL | F | 7 ans |
| Contrôle Sain RNAseq et qPCR | NA | Fibroblaste | F | 9 ans |
| Contrôle Sain RNAseq et qPCR | NA | Fibroblaste | M | 9 ans |
| Contrôle Sain RNAseq | NA | Fibroblaste | F | 39 ans |

| | | | | | |
|--------------------|------|----|-------------|---|---------|
| Contrôle RNAseq | Sain | NA | Fibroblaste | F | Inconnu |
| Contrôle qPCR | Sain | NA | Fibroblaste | M | 1 an |
| Contrôle qPCR | Sain | NA | Fibroblaste | F | 3 mois |
| Contrôle qPCR | Sain | NA | Fibroblaste | M | 7 ans |
| Contrôle qPCR | Sain | NA | Fibroblaste | F | 16 ans |
| Contrôle RNAseq | Sain | NA | LCL | F | 46 ans |
| Contrôle RNAseq | Sain | NA | LCL | F | 39 ans |
| Contrôle RNAseq | Sain | NA | LCL | F | Inconnu |

2.1.2 Séquençage d'ARN et qPCR

La qualité de l'ARN a ensuite été analysée par la plateforme de séquençage au Centre Intégré de Génomique Clinique Pédiatrique CHU Sainte-Justine et Génome Québec (CIGCP) avec le Bioanalyseur Agilent 2100. Les librairies ont été préparées avec des échantillons ayant un RIN (RNA Integrity Number) de plus de 8. Les librairies ont été préparées avec le kit Illumina TruSeq Stranded mRNA V2 (Illumina) ou le kit NEBNext Ultra Directional II (New England Biosystems). Nous étions intéressés principalement par les ARNs codants car ceux-ci sont les

mieux annotés et permettent une analyse fonctionnelle plus simple, et nous avons choisi ces préparations des librairies sur les conseils de la plateforme de Génomique du CHU Sainte Justine. Les librairies ont ensuite été séquencées sur Illumina HiSeq4000 ou NovaSeqS2 avec au moins 25 millions de reads de 100pb par échantillon.

Pour la quantification d'ARNm par RT-qPCR, l'ARN est isolé de la même façon, traité à la DNase, puis 3 µg d'ARN sont rétro-transcrits avec le kit qScript cDNA Supermix (Quantabio). L'ADN complémentaire obtenu est ensuite utilisé à une dilution finale de 1/200 et quantifié avec le kit de qPCR PowerUp SYBR Green Master Mix (Thermo).

2.1.3 ATACseq

2 millions de fibroblastes congelés dans du DMEM +10% FBS, 1% P/S, 1% Glutamax, 10% DMSO ont été envoyées à la compagnie Quick Biology (Pasadena, USA). Les cellules ont été homogénéisées, les noyaux isolés, et les librairies préparées in-vitro selon le protocole de la compagnie. Les librairies ont ensuite été séquencées par Illumina HiSeq4000, avec 80 millions de reads par échantillon en paired-end, soit 40 millions de fragments.

Tableau IV: Séquence des amorces utilisées pour la quantification par PCR en temps réel. Les amorces ont été générées manuellement à partir des séquences de cDNA de l'isoforme le plus long pour chaque gène, téléchargées depuis NCBI Refseq et visualisées sur Geneious version 7.0. La spécificité et le potentiel de dimérisation a été vérifié sur NCBI PrimerBlast.

| Amorce | Séquence |
|--------------------|---------------------------|
| SE_hABCC9_qPCR_F | TCGCTTCCTTTTGAGTCCTG |
| SE_hABCC9_qPCR_R | ATGTCCTCTGTTTCTGCGG |
| SE_hKCNJ8_qPCR_F | GCTCTTCGCTATCATGTGGT |
| SE_hKCNJ8_qPCR_R | GAAGAGAAAAGCAGAAGTGAAAGAC |
| SE_hSLC12a8_qPCR_F | TGTCTTTGGGGTTTTCTTCCC |
| SE_hSLC12a8_qPCR_R | AACAGCTGCCAGGGAGCC |
| SE_hSLC12a7_qPCR_F | TGCGTCGTGCTGTCCATCC |
| SE_hSLC12a7_qPCR_R | TTGACGCAGGCATCGAAGC |
| SE_hADR2A_qPCR_F | TCGTCATCATCGCCGTGTTT |
| SE_hADR2A_qPCR_R | AAGCCTTGCCGAAGTACCAG |
| SE_hARRB1_qPCR_F | AAGAAGGCCAGTCCAAATGG |
| SE_hARRB1_qPCR_R | GTCCACGAGGTCGATGTGG |
| SE_hKCNS1_qPCR_F | TTCCGCGTACTCAAGTTGG |
| SE_hKCNS1_qPCR_R | AACACTGACACACCCACAGC |
| SE_hGRK3_qPCR_F | TGCAGGAAAGCAGACACTGG |
| SE_hGRK3_qPCR_R | TTAAGGCTAATGTTTCTCCTTG |
| SE_hSMARCC2_qPCR_F | AGTATCTTACCTCTACCGCC |
| SE_hSMARCC2_qPCR_R | TAAGACCCCACTGTTCTAGG |

2.2. Analyse Bioinformatique

Tous les fichiers contenant le code informatique que j'ai écrit sont disponibles sur github.com/CampeauPipelines.

2.2.1 Séquençage de l'ARN

Les bibliothèques ont été analysées avec le pipeline bioinformatique Genpipes de MUGQIC (McGill University Genome Quebec and Innovation Centre)²⁴¹. En résumé, les fichiers BAM ont été retransformés en FASTQ avec le paquet Picard²⁴², les séquences ont été raccourcies avec Trimmomatic²⁴³, puis alignées sur le génome GRCh38 avec le paquet STAR²⁴⁴. Les séquences

dupliquées et mal alignées ont été enlevées avec Picard, puis les comptes par gène ont été déterminés avec le paquet Htseq²⁴⁵. Ensuite, l'analyse de l'expression différentielle a été réalisée avec le paquet DESeq2 sur R²⁴⁶. Les annotations GO (Gene Ontology) et leur enrichissement sont analysés avec le logiciel web GOrilla²⁴⁷ ou Panther²⁴⁸.

2.2.2 ATACseq

Les fichiers FASTQs ont été raccourcis avec Trimmomatic²⁴³, puis alignées avec Bowtie2 avec les paramètres additionnels --local et --mm²⁴⁹. Les fichiers ont ensuite été organisés et les duplicatas annotés et enlevés avec Samtools et Picard^{242,250}. Enfin, les pics ont été détectés avec MACS2 avec les paramètres --broad, --nomodel, --extsize 200, --nolambda, et l'analyse statistique réalisée avec Homer (findPeaks.pl et AnnotatePeaks.pl)^{251,252}.

2.2.3 Analyse du regroupement statistique des mutations

L'analyse a été effectuée sur le transcrit le plus long de chaque gène selon la méthode déjà publiée¹⁸⁶. Pour les mutations introniques dans les sites d'épissage, la mutation est annotée selon le nucléotide codant le plus proche.

3. Résultats

3.1 Des mutations dans le gène *TRRAP* sont associées à un syndrome neurodéveloppemental et induisent des changements d'expression génique

3.1.1 Mise en situation

Le gène *TRRAP* encode une protéine de près de 4000 acides-aminés qui est présent dans plusieurs complexes histone acétyl-transférase et interagit avec E2F et Myc pour activer leur activité transcriptionnelle. Nous avons identifié des mutations dans le gène *TRRAP* associées à un syndrome neurodéveloppemental, et avons montré un impact de certaines de ces mutations sur le profil d'expression de cellules provenant d'individus affectés.

Benjamin Cogné a organisé le recrutement des patients et leur regroupement et a, avec Eliane Beauregard-Lacroix, étudié le phénotype des individus affectés. J'ai cultivé les cellules des individus et extrait l'ARN, puis j'ai effectué les analyses bio-informatiques. Thomas Garcia a fait les RT-qPCR pour vérifier les gènes ciblés. Slavé Petrovski a réalisé l'image composite des photos d'individus affectés. Benjamin, Eliane, Justine et moi avons participé à la rédaction de l'article. Les autres auteurs sont des cliniciens qui ont participé au recrutement de patients et à la collecte des données. Le papier est publié : Cogné et al., *Am J Hum Gen* 2019²²⁶.

3.1.2 Article 1

Missense variants in the histone acetyltransferase complex component gene *TRRAP* cause autism and syndromic intellectual disability. Benjamin Cogné,^{1,2,80} Sophie Ehresmann,^{3,80} Eliane Beauregard-Lacroix,^{3,80} Justine Rousseau,³ Thomas Besnard,^{1,2} Thomas Garcia,³ Slavé Petrovski,^{4,5} Shiri Avni,⁶ Kirsty McWalter,⁷ Patrick R. Blackburn,^{8,9} Stephan Sanders,¹⁰ Kévin Uguen,^{11,12} Jacqueline Harris,^{13,14} Julie S. Cohen,¹³ Moira Blyth,¹⁵ Anna Lehman,¹⁶ Jonathan Berg,¹⁷ Mindy H. Li,¹⁸ Usha Kini,¹⁹ Shelagh Joss,²⁰ Charlotte von der Lippe,²¹ Christopher T. Gordon,^{22,23} Jennifer B. Humberson,²⁴ Laurie Robak,²⁵ Daryl A. Scott,^{25,26,27} Vernon R. Sutton,^{25,26,28} Cara M. Skraban,^{29,30} Jennifer J. Johnston,³¹ Annapurna Poduri,³² Magnus Nordenskjöld,^{33,34} Vandana Shashi,³⁵ Erica H. Gerkes,³⁶ Ernie M. Bongers,³⁷ Christian Gilissen,³⁷ Yuri A. Zarate,³⁸ Malin Kvarnung,^{33,34} Kevin P. Lally,³⁹ Peggy A. Kulch,⁴⁰ Brina Daniels,³⁸ Andres Hernandez-Garcia,²⁵ Nicholas Stong,⁴¹ Julie McGaughan,⁴² Kyle Retterer,⁷ Kristian Tveten,⁴³ Jennifer Sullivan,³⁵ Madeleine R. Geisheker,⁴⁴ Asbjorg Stray-Pedersen,⁴⁵ Jennifer M. Tarpinian,⁴⁶ Eric W. Klee,^{8,9,47,48} Julie C. Sapp,³¹ Jacob Zyskind,⁷ Øystein L. Holla,⁴³ Emma Bedoukian,⁴⁶ Francesca Filippini,^{22,23} Anne Guimier,^{22,23,49} Arnaud Picard,^{23,50} Øyvind L. Busk,⁴³ Jaya Punetha,²⁵ Rolph Pfundt,³⁷ Anna Lindstrand,^{33,34} Ann Nordgren,^{33,34} Fayth Kalb,⁵¹ Megha Desai,⁷ Ashley Harmon Ebanks,³⁹ Shalini N. Jhangiani,⁵² Tammie Dewan,¹⁶ Zeynep H. Coban Akdemir,²⁵ Aida Telegrafi,⁷ Elaine H. Zackai,^{29,30} Amber Begtrup,⁷ Xiaofei Song,²⁵ Annick Toutain,^{53,54} Ingrid M. Wentzensen,⁷ Sylvie Odent,^{55,56} Dominique Bonneau,^{57,58} Xénia Latypova,^{1,2} Wallid Deb,^{1,2} CAUSES Study,¹⁶ Sylvia Redon,^{11,12} Frédéric Bilan,^{59,60} Marine Legendre,^{59,60} Caitlin Troyer,²⁴ Kerri Whitlock,⁶¹ Oana Caluseriu,⁶¹ Marine I. Murphree,⁴⁷ Pavel N. Pichurin,⁴⁷ Katherine Agre,⁴⁷ Ralitza Gavriloova,^{47,62} Tuula Rinne,³⁷ Meredith Park,⁶³ Catherine Shain,⁶⁴ Erin L. Heinzen,⁴¹ Rui Xiao,^{25,65} Jeanne Amiel,^{22,23,49} Stanislas Lyonnet,^{22,23,49} Bertrand Isidor,^{1,2} Leslie G. Biesecker,³¹ Dan Lowenstein,⁶⁶ Jennifer

E. Posey,²⁵ Anne-Sophie Denommé-Pichon,^{57,58} Deciphering Developmental Disorder study,⁶⁷ Claude Férec,^{11,12} Xiang-Jiao Yang,^{68,69} Jill A. Rosenfeld,²⁵ Brigitte Gilbert-Dussardier,^{59,60} Séverine Audebert-Bellanger,⁷⁰ Richard Redon,² Holly A. F. Stessman,⁷¹ Christoffer Nellaker,^{72,73,74} Yaping Yang,^{25,65} James R. Lupski,^{25,26,52,75} David B. Goldstein,⁴¹ Evan E. Eichler,^{44,76} Francois Bolduc,^{61,77,78} Stéphane Bézieau,^{1,2} Sébastien Küry,^{1,2,81,*} Philippe M. Campeau^{3,79,81,*}

1. CHU Nantes, Service de Génétique Médicale, 9 quai Moncousu, 44093, France;
2. l'Institut du Thorax, INSERM, CNRS, UNIV Nantes, 44007 Nantes, France;
3. Centre Hospitalier Universitaire Sainte-Justine Research Centre, University of Montreal, Montreal, QC H3T 1C5, Canada;
4. Department of Medicine, The University of Melbourne, Melbourne, Australia;
5. AstraZeneca Centre for Genomics Research, Precision Medicine and Genomics, IMED Biotech Unit, AstraZeneca, Cambridge CB2 0AA, UK;
6. Visual Geometry Group, Dept. of Engineering Science, University of Oxford, Oxford OX1 3PJ, UK;
7. GeneDx, 207 Perry Parkway, Gaithersburg, Maryland 20877, USA;
8. Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, 55905, USA;
9. Center for Individualized Medicine, Health Sciences Research, Mayo Clinic, Rochester, MN 55905, USA;
10. Department of Psychiatry, UCSF Weill Institute for Neurosciences, University of California, San Francisco, San Francisco, California 94158, USA;
11. UMR1078 "Génétique, Génomique Fonctionnelle et Biotechnologies", Inserm, EFS, Université de Brest, ISBAM, 29200 Brest, France;
12. Laboratoire de Génétique Moléculaire et d'Histocompatibilité, CHRU Brest, 29200 Brest, France;
13. Division of Neurogenetics and Hugo W. Moser Research Institute, Kennedy Krieger Institute, Baltimore, MD 21205, USA;
14. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA;

15. Yorkshire Regional Genetics Service, Leeds Teaching Hospitals NHS Trust, Department of Clinical Genetics, Chapel Allerton Hospital, Chapeltown Road, Leeds LS7 4SA, UK;
16. Department of Medical Genetics, Children's and Women's Health Centre of British Columbia, Vancouver, BC V6H 3N1, Canada;
17. Molecular and Clinical Medicine, School of Medicine, University of Dundee, Ninewells Hospital & Medical School, Dundee DD1 9SY, UK;
18. Rush University Medical Center, Department of Pediatrics, Division of Genetics, Chicago, IL, 60612, USA;
19. Oxford Centre for Genomic Medicine, Oxford University Hospitals NHS Trust, Oxford OX3 7LE, UK;
20. West of Scotland Regional Genetics Service, Queen Elizabeth University Hospital, Glasgow G51 4TF, UK;
21. Department of Medical Genetics, St. Olav's Hospital, Trondheim University Hospital, 7006 Trondheim, Norway;
22. Laboratory of Embryology and Genetics of Human Malformations, Institut National de la Santé et de la Recherche Médicale (INSERM) UMR 1163, Institut Imagine, 75015 Paris, France;
23. Paris Descartes-Sorbonne Paris Cité University, Institut Imagine, 75015 Paris, France;
24. Division of Genetics, Department of Pediatrics, University of Virginia Children's Hospital, Charlottesville, Virginia 22903, USA;
25. Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA;
26. Texas Children's Hospital, Houston, TX 77030, USA;
27. Department of Molecular Physiology and Biophysics, Baylor College of Medicine, Houston, TX 77030, USA;
28. Baylor Genetics, Houston, TX 77030, USA;
29. Division of Human Genetics, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA;
30. Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA;
31. Medical Genomics and Metabolic Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892-4472, USA;
32. Division of Epilepsy and Clinical Neurophysiology and Epilepsy Genetics Program, Boston Children's Hospital, Harvard Medical School, Boston, MA 02115, USA;

33. Department of Molecular Medicine and Surgery, Center for Molecular Medicine, Karolinska Institutet, 17176 Stockholm, Sweden;
34. Department of Clinical Genetics, Karolinska University Hospital, 17176 Stockholm, Sweden;
35. Division of Medical Genetics, Department of Pediatrics, Duke University Medical Center, Durham, NC 27710, USA;
36. University of Groningen, University Medical Center Groningen, Department of Genetics, Groningen, 9700 RB, The Netherlands;
37. Department of Human Genetics, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, 6525 GA, the Netherlands;
38. Section of Genetics and Metabolism, University of Arkansas for Medical Sciences, Little Rock, Arkansas 72202, USA;
39. Department of Pediatric Surgery, The McGovern Medical School at the University of Texas Health Science Center at Houston, 6431 Fannin Street, Suite 5.258 Houston, Texas 77030 Children's Memorial Hermann Hospital, Houston, Texas;
40. Division of Genetics and Metabolism, Phoenix Children's Hospital, Phoenix, AZ 85006, USA;
41. Institute for Genomic Medicine, Columbia University, New York, NY 10032, USA;
42. Genetic Health Queensland, Royal Brisbane and Women's Hospital, Brisbane, Queensland and School of Medicine, The University of Queensland, Brisbane, Queensland 4029, Australia;
43. Department of Medical Genetics, Telemark Hospital Trust, 3710 Skien, Norway;
44. Department of Genome Sciences, University of Washington School of Medicine, Seattle, Washington 98195, USA;
45. Division of Pediatric and Adolescent Medicine, Oslo University Hospital - Rikshospitalet, Pb 4950 Nydalen, N-0424 Oslo, Norway;
46. Roberts Individualized Medical Genetics Center, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104, USA;
47. Department of Clinical Genomics, Mayo Clinic, Rochester, MN, 55905, USA;
48. Department of Biomedical Informatics, Mayo Clinic, Rochester, MN, 55905, USA;
49. Service de Génétique, Hôpital Necker-Enfants Malades, Assistance Publique - Hôpitaux de Paris (APHP), 75015 Paris, France;
50. Service de Chirurgie Maxillofaciale et Plastique, Centre de référence MAFACE, Hôpital Necker-Enfants Malades, Assistance Publique - Hôpitaux de Paris (APHP), 75015 Paris, France;

51. Division of Genetics, Birth Defects, and Metabolism, Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, IL 60611, USA;
52. Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX 77030, USA;
53. CHU Tours, Service de Génétique, 2 Boulevard Tonnellé, 37044 Tours, France;
54. INSERM U1253, IBrain, Université de Tours, 37032 Tours, France;
55. Service de Génétique Clinique, Centre Référence “Déficiences Intellectuelles de causes rares” (CRDI), Centre de référence anomalies du développement CLAD-Ouest, CHU Rennes, 35203 Rennes, France;
56. CNRS UMR 6290, Université de Rennes, 2 Avenue du Professeur Léon Bernard, 35043 Rennes, France;
57. CHU Angers, Département de Biochimie et Génétique, 49933 Angers Cedex 9, France;
58. MITOVASC, UMR CNRS 6015-INSERM 1083, Université d'Angers, 49933 Angers, France;
59. CHU Poitiers, Service de Génétique, BP577, 86021 Poitiers, France;
60. EA 3808, Université Poitiers, 86034, France;
61. Department of Medical Genetics, University of Alberta, Edmonton, AB T6G 2H7, Canada;
62. Department of Neurology, Mayo Clinic, Rochester, MN, 55905, USA;
63. Epilepsy Genetics Program, Department of Neurology, Boston Children's Hospital, Boston, Massachusetts 02115, USA;
64. Department of Neurology, Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA;
65. Baylor Genetics, Houston, TX, 77021, USA;
66. Department of Neurology, University of California, San Francisco, San Francisco, California 94143, USA;
67. Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton CB10 1SA, UK;
68. Rosalind & Morris Goodman Cancer Research Center and Department of Medicine, McGill University, Montreal, QC H3A 1A3, Canada;
69. Department of Biochemistry, McGill University and McGill University Health Center, Montreal, QC H3A 1A3, Canada;
70. CHRU Brest, Génétique médicale, 29609 Brest, France;

71. Department of Pharmacology, Creighton University Medical School, Omaha, NE 68178, USA;

72. Nuffield Department of Women's and Reproductive Health, University of Oxford, Oxford, UK;

73. Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford, UK;

74. Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Oxford OX3 7FZ, UK;

75. Department of Pediatrics, Baylor College of Medicine, Houston, TX 77030, USA;

76. Howard Hughes Medical Institute, Seattle, Washington 98195, USA;

77. Division of Pediatric Neurology, University of Alberta, Edmonton, Alberta, Canada;

78. Neuroscience and Mental Health Institute, University of Alberta, Edmonton, Alberta, Canada;

79. Department of Pediatrics, University of Montreal, Montreal, QC H3T1J4, Canada;

80. These authors contributed equally to this work.

81. These authors contributed equally to this work.

*Corresponding authors: p.campeau@umontreal.ca, sebastien.kury@chu-nantes.fr

Abstract

Acetylation of the lysine residues in histones and other DNA-binding proteins play a major role in regulation of eukaryotic gene expression. This process is controlled by histone acetyltransferases (HATs/KATs) found in multiprotein complexes that are recruited to chromatin by the scaffolding subunit TRansformation/tRanscription domain-Associated Protein (TRRAP). *TRRAP* is evolutionarily conserved and is among the top five genes intolerant to missense variation. Through an international collaboration, 17 distinct *de novo* or apparently *de novo* variants were identified in *TRRAP* in 24 individuals. A strong genotype-phenotype correlation was observed with two distinct clinical spectra. The first is a complex, multi-systemic syndrome characterized by a wide range of intellectual functioning, including a number of individuals with intellectual disability (ID) and markedly impaired basic life functions and associated with various malformations of the brain (i.e. cerebellar vermis hypoplasia and microcephaly), heart, kidneys, and genitourinary system. Individuals with this phenotype had missense variants clustering around the c.3127G>A p.(Ala1043Thr) variant identified in five individuals. The second spectrum manifested with autism spectrum disorder and/or ID and epilepsy. Facial dysmorphism was seen in both groups and included upslanted palpebral fissures, epicanthus, telecanthus, a wide nasal bridge and ridge, a broad and smooth philtrum and a thin upper lip. RNA sequencing analysis on affected individual-derived skin fibroblasts showed significant changes in expression of several genes implicated in neuronal function and ion transport. Thus, we describe here the clinical spectrum associated with *TRRAP* pathogenic missense variants and suggest a genotype-phenotype correlation useful for clinical evaluation of the pathogenicity of the variants.

Keywords:

TRRAP, histone acetylation, *de novo* variants, intellectual disability, congenital malformations, autism spectrum disorder

Introduction

Post-translational modifications of core histones including acetylation, methylation, phosphorylation, and ubiquitination directly alter DNA-histone and histone-histone interactions, thus influencing nucleosome dynamics^{1(p)}. Tight regulation of these marks is required by cells for proper gene transcription², DNA repair³, and DNA replication. One major activator of transcription is the acetylation of histone tails, which act by neutralizing the positive charges of lysine residues or by recruiting chromatin remodelers and transcription factors⁴. This tightly regulated process is performed by histone acetyltransferases (HAT) and reversed by histone deacetylases (HDAC). There are three major families of HAT: Gcn5-related N-acetyltransferase (GNAT), MOZ, YBF2/SAS3, SAS2, TIP60 (MYST), and EP300-CREBBP⁵. The activity and localization of most HATs like TIP60 or GCNL5 depend on a multiprotein assembly that contains the scaffolding protein TRansformation/tRanscription domain-Associated Protein (TRRAP).

TRRAP is a large protein of 3859 amino acids that is conserved from yeast to humans. It is an ataxia-telangiectasia mutated (ATM)-related member of the phosphatidylinositol (PI) 3-kinase-related kinase (PIKK) family⁶. Like other ATM-related members, it contains FAT (FRAP, ATM and TRRAP) and FATC (FRAP, ATM and TRRAP, C terminus) domains flanking a PI3/PI4-kinase domain. The kinase domain of TRRAP does not possess catalytic activity⁷, but is required for the proper recruitment of HAT complexes⁸. TRRAP has been shown to be involved in P53-, E2F- and c-MYC-dependent gene transcription and oncogenic transformation^{6,9,10}. As stressed in cancer studies, TRRAP plays an important role in cell cycle regulation. A recurrent somatic *TRRAP* variant c.2165C>T p.(Ser722Phe)¹¹ has been identified in melanoma and the oncogenic potential of TRRAP has been identified in glioblastoma multiforme¹², pancreatic adenocarcinoma¹³, and lymphoma¹⁰. Furthermore, *Trrap* knockout

leads to early embryonic lethality in mice through errors in the cell-cycle and a failure to arrest at the mitotic checkpoint¹⁴. In mouse embryonic stem cells (ESCs), *Trrap* is indispensable for self-renewal as well as correct differentiation¹⁵, suggesting an essential role in embryonic development and morphogenesis. Moreover, brain-specific *Trrap* knockout in mice leads to premature differentiation of neural progenitors and abnormal brain development through a decrease in the expression of cell-cycle regulators resulting in brain atrophy and microcephaly¹⁶. *TRRAP* has previously been associated with neuropsychiatric disorders such as schizophrenia in few patients^{17–20}. We herein provide data showing that *TRRAP* pathogenic variants are associated with a variable neurodevelopmental disorder.

Results:

Through an international collaboration and aided by the web-based tool GeneMatcher²¹, we identified 17 distinct missense variants in *TRRAP* with strong clinical and/or molecular evidence for pathogenicity in 24 individuals with neurodevelopmental disorders (**Table 1, Figure 8A**). These variants were identified either by trio or solo exome sequencing (ES) from research and clinical cohorts. All affected individuals or their guardians gave appropriate consent for research procedures. Methods are described in **Table S1**.

These 17 variants were absent from ExAC and gnomAD²², and were found *de novo* or apparently *de novo* (maternity and paternity not checked) in all individuals, except for two sisters with a variant inherited from a mother with low-level mosaicism (**Figure S1**) and an individual whose father was unavailable but whose paternal grandparents did not carry the variant. Three variants were recurrently observed: p.(Ala1043Thr) identified in five individuals and p.(Glu1106Lys) and p.(Gly1883Arg) identified in two individuals each. All the variants were predicted to be deleterious by CADD²³ (scaled C-scores over 20) and variously predicted pathogenic by SIFT²⁴ and PolyPhen-2 HVAR²⁵. As shown in **figure 8A**, the 17 variants seen in our subjects had significantly increased CADD scores compared to singleton missense variants reported in gnomAD.

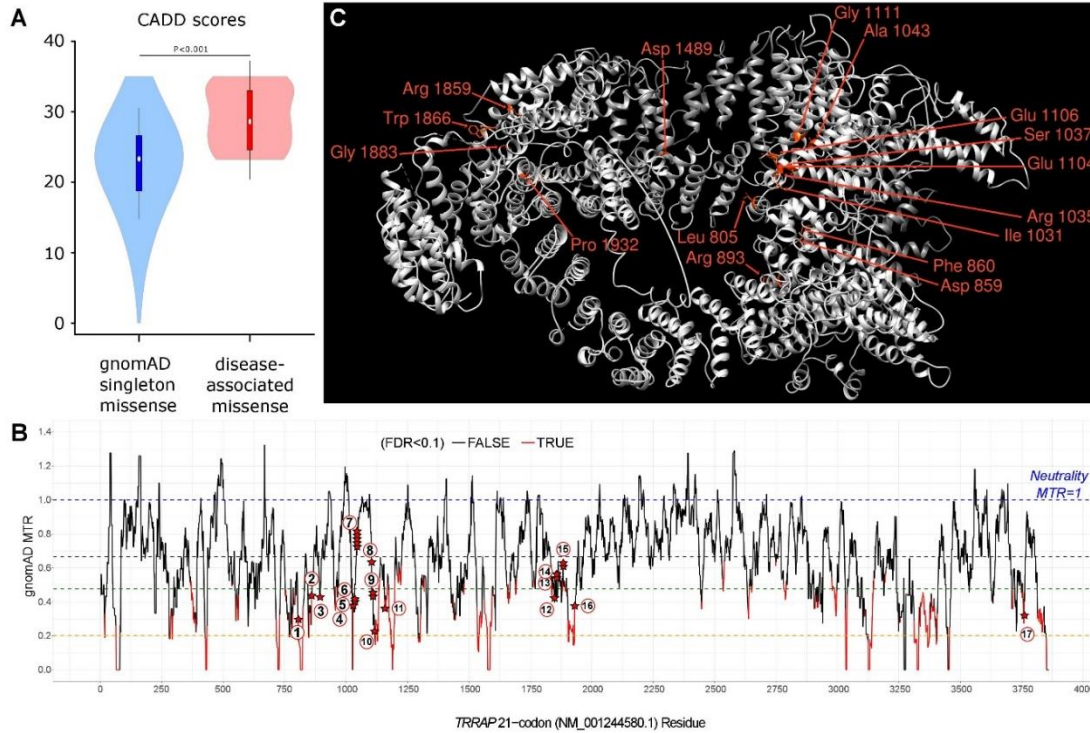


Figure 9: *TRRAP* sequence is intolerant to missense variants.

(A) CADD scores of the 17 variants identified in affected individuals compared to gnomAD singleton missense variants. The individuals' variants were compared to variants seen once in gnomAD, in order to avoid CADD training circularity. (B) *TRRAP* Missense Tolerance Ratio (MTR) plot. (C) Localization of the *TRRAP* residues affected in the individuals in 3D protein models, including 14 out of 17 likely pathogenic variants and 2 out of 6 additional variants of unknown significance. The representation of the structure of human *TRRAP* (NP_001231509.1) was predicted by PHYRE2 Protein Fold Recognition Server by comparison to its *Saccharomyces cerevisiae* ortholog, according to the cryo-EM structure of the SAGA and NuA4 coactivator subunit Tra1 present in the protein data bank (PDB ID: 5OJS). Variants in regions non-homologous to Tra1 are not represented. Structure representation was made with UCSF Chimera.

The 17 variants all occurred at residues conserved among vertebrates (**Figure 8B**), and in regions depleted in missense variants in gnomAD. Indeed, assessing Missense Tolerance Ratios for *TRRAP*, we observed that most of the 17 variants were in regions intolerant to missense variants (**Figure 9B**). Nine out of the 17 variants occurred at highly mutable CpG sites, including one within the codon that leads to the recurrent p.(Ala1043Thr) variant observed in five individuals. Six missense variants with lesser evidence for pathogenicity were found in another six unrelated individuals (Individuals 25 to 30 in **Table S1**). These variants may be deleterious but were not clearly pathogenic, since either the inheritance pattern could not be determined, the variant was present in gnomAD or lead to another missense change at the same residue as a variant reported in gnomAD, or were located in a less conserved region of *TRRAP* (**Table S2**).

Given the number of *de novo* variants identified, the enrichment for *TRRAP de novo* variants in our study was calculated as ($p=4.2 \times 10^{-6}$) based on denovolyzer²⁶. Nevertheless, the current number of 22 detected *de novo* variants in *TRRAP* is not genome-wide significant ($p=0.08$) after correcting for: a) ~19,000 protein-coding genes, b) 22,898 trios studied, and c) the underlying mutability of the full-length protein-coding *TRRAP* transcript. However, this statistical calculation does not take into account the spatial distribution of the variants. Indeed, three-dimensional modelling of human TRRAP structure inferred from the orthologous *Saccharomyces cerevisiae* protein Tra1 (**Figure 9C**) suggested a clustering of the variants in different regions of TRRAP. The most important clustering was observed for thirteen variants between codons 1031 and 1159. Interestingly, when visualized in 3D, these variants localized near one another (**Figure 8C**), revealing a domain of TRRAP with a potentially novel specific function, although not yet characterized. We performed a statistical clustering analysis comparing the mean distance in observed variants to ten million permutations of random

variants, as previously described²⁷. This analysis revealed a significant clustering of variants along the primary sequence of *TRRAP* (p value = 9×10^{-8}), suggesting a model where specific domains are affected, haploinsufficiency being unlikely, at least for clustering variants.

Among the 24 individuals who carried pathogenic variants, 19 presented with facial dysmorphisms. Recurrent features that were noted among these individuals included upslanted palpebral fissures, epicanthus, telecanthus, a wide nasal bridge and ridge, a broad and smooth philtrum and a thin upper lip (**Figure 10**). We performed a computer-assisted facial gestalt visualisation^{28,29}, which highlights several of these features, particularly for individuals with variants clustering with the recurrent p.(Ala1043Thr) variant (**Figure 10R**). All the individuals had developmental delay, with a highly variable severity of intellectual disability (ID). Whereas most individuals had apparent ID with markedly impaired basic life functions, some of them presented with mild ID or even no cognitive deficits (**Table 2 and Table S3**). Peripheral neuropathy was also noted; it was severe in one individual and consisted of lower limb hyperreflexia in five other individuals.



Figure 10: Photographs of individuals with *TRRAP* variants.

In addition to alteration in cerebral function, some individuals showed brain, cerebellum, heart, kidney or urogenital malformations. We observed a strong genotype-phenotype

correlation (**Figure 7A, Table 2**) with the highest incidence of malformations being seen in thirteen individuals whose variants cluster in the region of the predicted protein from codons 1031 to 1159: c.3093T>G p.(Ile1031Met), c.3104G>A p.(Arg1035Gln), c.3111C>A p.(Ser1037Arg), c.3127G>A p.(Ala1043Thr), c.3311A>G p.(Glu1104Gly), c.3316G>A p.(Glu1106Lys), c.3331G>T p.(Gly1111Trp), and c.3475G>A p.(Gly1159Arg). In contrast, individuals with variants residing outside of this region had less malformation and presented mainly with autism spectrum disorder (ASD) and/or ID, sometimes associated with epilepsy. Variants in these individuals were more dispersed along the protein, although some apparently aggregated in another region, including c.5575C>T p.(Arg1859Cys), c.5596T>A p.(Trp1866Arg), c.5598G>T p.(Trp1866Cys), c.5647G>A p.(Gly1883Arg), and c.5795C>T p.(Pro1932Leu).

Thirteen individuals with variants in the codon 1031-1159 region had global developmental delay and apparent ID, ranging from speech delay and learning difficulties to markedly impaired basic life functions (**Table 2 and Table S3**). The last available occipitofrontal circumference measurements revealed microcephaly (ranging from -2.8 to -5 standard deviations (SDs)) in 46% (6/13) of individuals. Cerebral magnetic resonance imaging (MRI) had been performed in 10/13 individuals and showed structural brain anomalies in seven of them (70%), including cerebellar vermis hypoplasia (6/10), ventricular enlargement (3/10), cortical atrophy (2/10), brainstem atrophy (2/10), polymicrogyria (1/10), focal gliosis (1/10), delayed myelination (1/10), and corpus callosum hypoplasia (1/10). Neurological examination revealed hypotonia in 31% (4/13) of individuals. Only one individual was reported with epilepsy. Seven individuals (54%) were reported to have feeding difficulties and were fed exclusively by gastrostomy tube. Among the 10 individuals who were examined by echocardiography, 70% (7/10) had abnormal results with 50% (5/10) having ventricular septal

defect, 30% (3/10) patent ductus arteriosus, 30% (3/10) patent foramen ovale, 20% (2/10) pulmonary hypertension, and 20% (2/10) aortic coarctation. Abdominal ultrasound revealed anomalies in 70% (7/10) of individuals in which it was performed. Abnormal renal morphology, namely multicystic dysplastic kidney, hydronephrosis, duplicate kidney and/or small kidney, was described in 60% (6/10) and vesicoureteral reflux was also observed in 30% (3/10) of these individuals. Individual 15 presented with a large left-sided posterolateral congenital diaphragmatic hernia (Table S3). Hernias of the abdominal wall were also found in 23% (3/13) of individuals and included an umbilical hernia, an omphalocele, and an inguinal hernia. Three males (3/6; 50%) had anomalies of external genitalia, including microphallus, hypoplastic scrotum, and cryptorchidism, while two females (2/7; 29%) had a duplicated vagina and/or uterus. Other observed anomalies included dysplastic nails (8/13; 62%), cleft lip/palate (5/13; 38%), clinodactyly of the 5th finger (4/13; 31%), laryngo/tracheomalacia (3/13), accessory nipple (3/13; 23%), bilateral cutaneous syndactyly of the second and third toe (2/13; 15%) and anomalies of the lacrimal glands (1/13; 8%; see also below in individuals 1 and 19). Four individuals (4/13; 31%) had visual impairment and three (3/13; 23%) had hearing impairment which has associated with inner ear malformations in 2 cases. Recurrent infections, mainly respiratory and urinary tract infections, affected 3/13 (23%) individuals. Individual 9 died at 12 years of age in the context of multiple co-morbidities including: renal failure with acute fluid fluctuations, tracheostomy for severely obstructive laryngotracheomalacia, intermittent supraventricular tachycardia, arterial insufficiency, and polyendocrinopathy (insulin dependent diabetes, adrenal insufficiency, and hypothyroidism).

Among individuals with variants falling outside of the 1031-1159 region, 5/11 (45%) were diagnosed with autism spectrum disorder (ASD), and another three individuals (3/11; 27%) had some findings of ASD but no formal diagnosis. Eight of 11 (73%) had developmental delay

and mild to severe intellectual disability and three had speech delay but whose IQ was measured above 70, including two that were in the normal range. Four individuals (4/11; 36%) had various types of epilepsy, namely absence and tonic-clonic seizures, or Lennox-Gastaut syndrome. The age of seizure onset ranged from 2 to 10 years old. Malformations were infrequent in this group overall, although individual 2 had microcephaly and heart malformations, individual 1 had lacrimal duct aplasia, individual 19 had lacrimal duct aplasia and optic disc colobomas and individual 21 had a postaxial polydactyly of one hand.

TRRAP-associated chromatin remodeling complexes are generally associated with gene activation³⁰, which is consistent with their HAT activity. Nevertheless, the NuA4 complex has been shown to have a gene repression activity necessary for embryonic stem cell pluripotency^{31,32} which seems to be independent from its lysine acetyltransferase activity³³. To test the hypothesis that *TRRAP* variants alter gene expression, we obtained skin fibroblasts from two individuals (individual 1 with p.(Leu805Phe) and individual 19 with p.(Trp1866Cys)) and performed next-generation sequencing with technical replicates of RNA (i.e. separately prepared libraries from the same samples). The RNA library preparation and sequencing as well as bioinformatics analysis methods can be found in the Supplementary notes. We found that, in comparison to two typically developing individuals (controls), both individuals with *TRRAP* variants had remarkably different gene expression patterns (**Figure S2A**). Interestingly, most differentially expressed genes (DEGs) analyzed with DESeq2 were upregulated in affected individuals compared to controls (**Figure S2B**). Moreover, the individual with p.(Leu805Phe) had 619 differentially expressed genes (DEGs), with a Log₂ Fold Change (Log₂FC) higher than 2 or lower than -2, and a p-value adjusted for 10%FDR lower than 0.01 (padj) (**supplementary notes, Table S5**).

To identify genes with significant expression differences, we performed differential gene expression analysis between the two individuals with *TRRAP* variants (combined as biological replicates) and two unaffected controls. Gene ontology (GO) enrichment analysis of these genes using the GOrilla web application indicates an enrichment for the adrenergic receptor signalling pathway, genes important for neurological function, and potassium and ATP-sensitive ion transporters (**Figure S2B, Supplementary notes, Table S5**). The two individuals tested carried variants outside the cluster associated with the more syndromic ID; assuming distinct effects on gene regulation, it will be worth comparing gene expression between the two groups. Finally, we cannot exclude that the transcriptome alteration may be caused by a mechanism other than impaired HAT activity, since it has been shown that TRRAP has direct interactions with different partners not related to HAT complex. Thus, we highlighted candidate pathways that may be useful to uncover the pathomechanism of *TRRAP* variants in future studies.

Discussion:

TRRAP acts as a scaffold in HAT complexes. Although it has not a direct role in acetylation, we hypothesize that pathogenic effect of variants may be due to dysregulation of acetylation, a major process that has been associated with several neurodevelopmental disorders³⁴. Pathogenic variants of *KAT6B* (MIM: 605880) cause both Say-Barber-Biesecker-Young-Simpson syndrome (SBBYSS [MIM: 603736])^{35–37} and genitopatellar syndrome (GPS [MIM: 606170])^{38,39} and pathogenic variants in *KAT6A* and *BRPF1* mutations have also been associated with a neurodevelopmental disorder^{40,41,42(p1)}. Rubinstein-Taybi syndrome (MIM: 180849 and 613684) is associated with variants in genes encoding histone acetyltransferase complexes, namely *CREBBP* and *EP300*^{43–46}. In addition to cognitive impairment, abnormal histone acetylation can also result in behavioral disorders, as evidenced by the associations

found between alterations in several lysine acetyltransferase (KAT) and lysine deacetylase (KDAC) genes including *BRD1*, *HDAC4*, *HDAC6*, and *HDAC9* and non-syndromic ASD and/or schizophrenia^{34,47,48,49(p9),50}.

Variants in *TRRAP* were associated with neuropsychiatric disorders including childhood disintegrative disorder¹⁷, schizophrenia^{18,19} and ASD²⁰. The ASD report included individuals 18 and 19 with p.(Trp1866Arg) and p.(Trp1866Cys), respectively. We thus confirmed the association with ASD, and provide evidence that it can be found either isolated, or associated with intellectual disability. Based on the ExAC dataset without studies on neuropsychiatric disorders, *TRRAP* is in the top five genes most intolerant to missense variants in the human genome, with a missense z-score of 10.1 (Ref. ²²). While this study includes only the first 24 identified individuals, a strength of the study is that it was primarily ascertained by sequencing, reducing phenotypic ascertainment bias. Given the highly constrained region of the observed variants coupled with the population constraint and evolutionary conservation, we hypothesize that variants outside of these regions are likely to be associated with prenatal lethality, although we cannot exclude that milder phenotypes may be underrepresented in current exomes datasets. It is worth noting that we exclusively identified missense variants in the affected individuals. Given the loss of function (LOF) intolerance of *TRRAP* in ExAC (pLI = 1.00), we would expect to identify at least some LOF variants if haploinsufficiency of *TRRAP* was the causal mechanism. In DECIPHER (accessed 2018-05-14), no small or intragenic deletions involving *TRRAP* have been identified. Thus, taking into account the significant clustering, it suggests that missense variants may act either as gain of function or dominant-negative and that haploinsufficiency of *TRRAP* is likely to be prenatally lethal, although we cannot exclude a LOF effect of non-clustering variants associated with a milder phenotype.

TRRAP participates in embryonic development, as demonstrated by its binding with proteins regulating signaling pathways such as Notch in fruit fly⁵¹, Ras in *C. elegans*⁵², or Wnt by influencing beta-catenine ubiquitinylation⁵³. We suspect therefore that TRRAP variants, more especially those falling within the 1031-1159 region, perturb the interactions with at least one of these developmental signaling pathways, which would explain the multiple malformation observed in about half of the affected individuals.

In yeast, a series of ~100 codon deletion mutants in the ortholog *tra1*, showed reduced or complete loss of viability⁵⁴. Most deletions impaired coactivator complex assembly, notably the ones encompassing the homologous 1031-1159 cluster (mutants $\Delta 13$ - $\Delta 14$), as well as the regions homologous to those containing variants p.(Leu805Phe), p.(Phe860Leu), and p.(Arg893Leu) (mutants $\Delta 11$ - $\Delta 12$) and the p.(Arg3757Gln) variant (mutant $\Delta 39$). In contrast, mutants $\Delta 21$ - $\Delta 22$ encompassing the region homologous to the cluster associated with less malformations (codons 1859-1932) were viable and may help explain the milder clinical phenotype associated with variants within this cluster. In mice, *Trrap* knock-out leads to early embryonic lethality¹⁴ and a neural cell-specific conditional *Trrap* knockout line¹⁶ revealed premature differentiation of neural progenitors, with depletion of progenitor pools and a significant reduction in cortical thickness. These mice exhibited striking microcephaly, in agreement with what we observed in half of the individuals in our study cohort, primarily those with variants in the 1031-1159 cluster.

In summary, we report evidence that variants in *TRRAP* are associated with a pleiotropic neurodevelopmental syndrome with a potential genotype-phenotype correlation. Our functional data highlight an enrichment of genes related to neuronal function and ion transport that could underline the pathophysiology of the disease. Future *in vitro* and *in vivo* studies on variants

inside and outside the main cluster will be required to dissect which gene expression changes are connected to which *TRRAP*-related specific phenotypes.

Acknowledgments

We would like to thank all families for participating in this study. We acknowledge HUGODIMS consortium, which was supported by a grant from the French Ministry of Health and from the Health Regional Agency from Poitou-Charentes (HUGODIMS, 2013, RC14_0107). We are grateful to Frédérique Allaire from the Health Regional Agency of Poitou-Charentes for supporting this project. We acknowledge Léa Ferrand and Emilie Le Blanc's assistance for grant and data management. We would like to thank the members of the Canadian Center for Computational Genomics and the McGill University and Génome Québec Innovation Center for their help in bioinformatics analysis. This work was also supported by funds from the National Institute of Neurological Disorders and Stroke (The Epilepsy Phenome/Genome Project NS053998; Epi4K NS077364, NS077274, NS077303, and NS077276) to D.L. and D.B.G. and from the National Institutes of Health/Eunice Kennedy Shriver National Institute of Child Health and Human Development grant (HD064667) to D.A.S. This work was supported in part by NINDS R35 NS105078 to JRL and jointly funded National Human Genome Research Institute (NHGRI), and National Heart, Lung, and Blood Institute (NHLBI) grant to the Baylor-Hopkins Center for Mendelian Genomics (UM1 HG006542). JEP was supported by NHGRI K08 HG008986. VS and JS were supported by the Duke Genome Sequencing Clinic grant. We thank the CIHR and FRSQ for clinician-scientist awards to PMC. We would also like to thank the Mayo Clinic Center for Individualized Medicine (CIM) for supporting this research through the CIM Investigative and Functional Genomics Program. E.E.E. is an investigator of the Howard Hughes Medical Institute. The CAUSES Study is funded by Mining for Miracles,

British Columbia Children's Hospital Foundation, and Genome British Columbia. CAUSES Study investigators include Shelin Adam, Christele Du Souich, Alison Elliott, Anna Lehman, Jill Mwenifumbo, Tanya Nelson, Clara Van Karnebeek, and Jan Friedman. LGB, JJJ, and JCS were supported by the intramural research program of the National Human Genome Research Institute, grant HG200328 12. We acknowledge also the DECIPHER Consortium which contributed to the exchange of genetic and clinical data between the teams. This work was supported, in part, by US National Institute of Mental Health grant R01MH101221 to E.E.E.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflicts of interest

E.E.E. is on the scientific advisory board (SAB) of DNAnexus, Inc. The Department of Molecular and Human Genetics at Baylor College of Medicine receives revenue for clinical genetic testing completed at Baylor Genetics laboratory. KMc, KR, JZ, MD, AT, AB, and IMW are employees of GeneDx, Inc. Dr. Goldstein is Founder and holds equity in Pairnomix and Praxis Therapeutics. Dr. Goldstein is not aware of any overlap with Pairnomix or Praxis Therapeutics.

Web Resources

ExAC Browser, <http://exac.broadinstitute.org/>

gnomAD, <http://gnomad.broadinstitute.org/>

Ensembl VEP, http://grch37.ensembl.org/Homo_sapiens/Tools/VEP

GenBank, <http://www.ncbi.nlm.nih.gov/genbank/>

OMIM, <http://www.omim.org/>

Phyre2, <http://www.sbg.bio.ic.ac.uk/phyre2/>

UniProt, <http://www.uniprot.org/uniprot/>

DECIPHER, <https://decipher.sanger.ac.uk/>

Tables

| cDNA | Protein | Inheritance | CpG | gnomAD | CADD score (v1.3) | SIFT | PolyPhen2 HVAR | Number of individuals |
|------------|----------------|---------------------------|-----|--------|-------------------|-------------------|--------------------------|-----------------------|
| c.2413C>T | p.(Leu805Phe) | <i>de novo</i> | no | absent | 28.2 | deleterious(0) | probably_damaging(0.998) | 1 |
| c.2580C>G | p.(Phe860Leu) | <i>de novo</i> | no | absent | 27.6 | deleterious(0.03) | possibly_damaging(0.867) | 1 |
| c.2678G>T | p.(Arg893Leu) | apparently <i>de novo</i> | yes | absent | 34 | deleterious(0) | probably_damaging(0.986) | 1 |
| c.3093T>G | p.(Ile1031Met) | <i>de novo</i> | no | absent | 23.4 | deleterious(0.02) | benign(0.308) | 1 |
| c.3104G>A | p.(Arg1035Gln) | <i>de novo</i> | yes | absent | 23.9 | tolerated(0.09) | benign(0.404) | 1 |
| c.3111C>A | p.(Ser1037Arg) | <i>de novo</i> | yes | absent | 23.7 | tolerated(0.14) | possibly_damaging(0.656) | 1 |
| c.3127G>A | p.(Ala1043Thr) | <i>de novo</i> | yes | absent | 23.2 | tolerated(0.27) | benign(0.066) | 5 |
| c.3311A>G | p.(Glu1104Gly) | <i>de novo</i> | no | absent | 24.6 | deleterious(0.04) | probably_damaging(0.91) | 1 |
| c.3316G>A | p.(Glu1106Lys) | <i>de novo</i> * | no | absent | 27.7 | deleterious(0) | possibly_damaging(0.816) | 2 |
| c.3331G>T | p.(Gly1111Trp) | apparently <i>de novo</i> | yes | absent | 34 | deleterious(0) | probably_damaging(0.999) | 1 |
| c.3475G>A | p.(Gly1159Arg) | <i>de novo</i> | no | absent | 33 | deleterious(0) | probably_damaging(0.999) | 1 |
| c.5575C>T | p.(Arg1859Cys) | <i>de novo</i> | yes | absent | 34 | deleterious(0) | probably_damaging(0.997) | 1 |
| c.5596T>A | p.(Trp1866Arg) | <i>de novo</i> | no | absent | 28.7 | deleterious(0) | probably_damaging(0.999) | 1 |
| c.5598G>T | p.(Trp1866Cys) | <i>de novo</i> | no | absent | 33 | deleterious(0) | probably_damaging(0.999) | 1 |
| c.5647G>A | p.(Gly1883Arg) | <i>de novo</i> | yes | absent | 33 | deleterious(0) | probably_damaging(1) | 2 |
| c.5795C>T | p.(Pro1932Leu) | germline mosaicism | yes | absent | 35 | deleterious(0) | probably_damaging(0.997) | 2 |
| c.11270G>A | p.(Arg3757Gln) | <i>de novo</i> | yes | absent | 28.6 | deleterious(0.01) | benign(0.269) | 1 |

Table 1. *De novo* variants in *TRRAP* identified in 24 individuals. RefSeq transcript used for *TRRAP* is NM_001244580.1. Apparently *de novo* was mentioned when paternity and maternity were not checked. *For one individual with p.(Glu1106Lys), father was unavailable, paternal grandparents were tested and did not carry the variant.

| Symptoms | All individuals | Cluster 1031-1159 | Variants outside the cluster |
|----------------------------|-----------------|-------------------|------------------------------|
| Global developmental delay | 24/24 - 100% | 13/13 - 100% | 11/11 - 100% |
| Intellectual disability | 17/20 - 85% | 11/11 - 100% | 6/9 - 67% |
| Facial dysmorphisms | 19/24 - 79% | 11/13 - 85% | 8/11 - 73% |
| Autism Spectrum Disorder | 5/24 - 21% | 0/13 - 0% | 5/11 - 45% |
| Microcephaly (<-2.5SD) | 7/24 - 29% | 6/13 - 46% | 1/11 - 9% |
| Short stature | 7/23 - 30% | 4/12 - 33% | 3/11 - 27% |
| Hypotonia | 8/24 - 33% | 4/13 - 31% | 4/11 - 36% |
| Feeding difficulties | 8/24 - 33% | 7/13 - 54% | 1/11 - 9% |
| Seizures | 5/24 - 21% | 1/13 - 8% | 4/11 - 36% |
| Cleft lip/palate | 5/24 - 21% | 5/13 - 38% | 0/11 - 0% |
| Cerebellar hypoplasia | 6/18 - 33% | 6/11 - 55% | 0/7 - 0% |
| Cerebral abnormalities | 6/18 - 33% | 6/11 - 55% | 0/7 - 0% |
| Cardiac malformations | 10/15 - 66% | 9/12 - 75% | 1/3 - 33% |
| Renal malformations | 5/17 - 29% | 5/13 - 38% | 0/4 - 0% |

| | | | |
|--------------------------|------------|------------|------------|
| Genital malformations | 5/24 - 21% | 5/13 - 38% | 0/11 - 0% |
| Hearing impairment | 3/24 - 12% | 3/13 - 23% | 0/11 - 0% |
| Visual impairment | 4/24 - 17% | 3/13 - 23% | 1/11 - 9% |
| Scoliosis | 3/24 - 12% | 3/13 - 23% | 0/11 - 0% |
| Dysplastic nails | 8/24 - 33% | 8/13 - 62% | 0/11 - 0% |
| Lower limb hyperreflexia | 5/24 - 21% | 1/13 - 8% | 4/11 - 36% |
| Lacrimal duct aplasia | 3/24 - 12% | 1/13 - 8% | 2/11 - 18% |
| Accessory nipple | 4/24 - 17% | 3/13 - 23% | 1/11 - 9% |

Table 2: Clinical description of individuals with variants inside or outside the 1031-1159 cluster.

REFERENCES

1. Bowman GD, Poirier MG. Post-translational modifications of histones that influence nucleosome dynamics. *Chem Rev.* 2015;115(6):2274-2295. doi:10.1021/cr500350x
2. Venkatesh S, Workman JL. Histone exchange, chromatin structure and the regulation of transcription. *Nat Rev Mol Cell Biol.* 2015;16(3):178-189. doi:10.1038/nrm3941
3. Hunt CR, Ramnarain D, Horikoshi N, et al. Histone Modifications and DNA Double-Strand Break Repair after Exposure to Ionizing Radiations. *Radiat Res.* 2013;179(4):383-392. doi:10.1667/RR3308.2
4. Legube G, Trouche D. Regulating histone acetyltransferases and deacetylases. *EMBO Rep.* 2003;4(10):944-947. doi:10.1038/sj.embor.embor941
5. Berndsen CE, Denu JM. Catalysis and substrate selection by histone/protein lysine acetyltransferases. *Curr Opin Struct Biol.* 2008;18(6):682-689. doi:10.1016/j.sbi.2008.11.004
6. McMahon SB, Van Buskirk HA, Dugan KA, Copeland TD, Cole MD. The novel ATM-related protein TRRAP is an essential cofactor for the c-Myc and E2F oncoproteins. *Cell.* 1998;94(3):363-374.
7. Vassilev A, Yamauchi J, Kotani T, et al. The 400 kDa subunit of the PCAF histone acetylase complex belongs to the ATM superfamily. *Mol Cell.* 1998;2(6):869-875.
8. Park J, Kunjibettu S, McMahon SB, Cole MD. The ATM-related domain of TRRAP is required for histone acetyltransferase recruitment and Myc-dependent oncogenesis. *Genes Dev.* 2001;15(13):1619-1624. doi:10.1101/gad.900101
9. Zhao L-J, Loewenstein PM, Green M. Enhanced MYC association with the NuA4 histone acetyltransferase complex mediated by the adenovirus E1A N-terminal domain activates a subset of MYC target genes highly expressed in cancer cells. *Genes Cancer.* 2017;8(11-12):752-761. doi:10.18632/genesandcancer.160
10. Jethwa A, Słabicki M, Hüllelin J, et al. TRRAP is essential for regulating the accumulation of mutant and wild-type p53 in lymphoma. *Blood.* April 2018. doi:10.1182/blood-2017-09-806679
11. Wei X, Walia V, Lin JC, et al. Exome sequencing identifies GRIN2A as frequently mutated in melanoma. *Nat Genet.* 2011;43(5):442-446. doi:10.1038/ng.810
12. Wurdak H, Zhu S, Romero A, et al. An RNAi screen identifies TRRAP as a regulator of brain tumor-initiating cell differentiation. *Cell Stem Cell.* 2010;6(1):37-47. doi:10.1016/j.stem.2009.11.002
13. Loukopoulos P, Shibata T, Katoh H, et al. Genome-wide array-based comparative genomic hybridization analysis of pancreatic adenocarcinoma: identification of genetic

- indicators that predict patient outcome. *Cancer Sci.* 2007;98(3):392-400. doi:10.1111/j.1349-7006.2007.00395.x
14. Herceg Z, Hulla W, Gell D, et al. Disruption of Trtrap causes early embryonic lethality and defects in cell cycle progression. *Nat Genet.* 2001;29(2):206-211. doi:10.1038/ng725
 15. Sawan C, Hernandez-Vargas H, Murr R, et al. Histone acetyltransferase cofactor Trtrap maintains self-renewal and restricts differentiation of embryonic stem cells. *Stem Cells.* 2013;31(5):979-991. doi:10.1002/stem.1341
 16. Tapias A, Zhou Z-W, Shi Y, et al. Trtrap-dependent histone acetylation specifically regulates cell-cycle gene transcription to control neural progenitor fate decisions. *Cell Stem Cell.* 2014;14(5):632-643. doi:10.1016/j.stem.2014.04.001
 17. Gupta AR, Westphal A, Yang DYJ, et al. Neurogenetic analysis of childhood disintegrative disorder. *Mol Autism.* 2017;8:19. doi:10.1186/s13229-017-0133-0
 18. Xu B, Ionita-Laza I, Roos JL, et al. De novo gene mutations highlight patterns of genetic and neural complexity in schizophrenia. *Nat Genet.* 2012;44(12):1365-1369. doi:10.1038/ng.2446
 19. Takata A, Xu B, Ionita-Laza I, Roos JL, Gogos JA, Karayiorgou M. Loss-of-Function Variants in Schizophrenia Risk and SETD1A as a Candidate Susceptibility Gene. *Neuron.* 2014;82(4):773-780. doi:10.1016/j.neuron.2014.04.043
 20. Geisheker MR, Heymann G, Wang T, et al. Hotspots of missense mutation identify neurodevelopmental disorder genes and functional domains. *Nat Neurosci.* 2017;20(8):1043-1051. doi:10.1038/nn.4589
 21. Sobreira N, Schiettecatte F, Valle D, Hamosh A. GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. *Hum Mutat.* 2015;36(10):928-930. doi:10.1002/humu.22844
 22. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature.* 2016;536(7616):285-291. doi:10.1038/nature19057
 23. Kircher M, Witten DM, Jain P, O’Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet.* 2014;46(3):310-315. doi:10.1038/ng.2892
 24. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc.* 2009;4(7):1073-1081. doi:10.1038/nprot.2009.86
 25. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods.* 2010;7(4):248-249. doi:10.1038/nmeth0410-248
 26. Ware JS, Samocha KE, Homsy J, Daly MJ. Interpreting de novo variation in human disease using denovolyzeR. *Curr Protoc Hum Genet.* 2015;87:7.25.1-7.25.15. doi:10.1002/0471142905.hg0725s87

27. Lelieveld SH, Wiel L, Venselaar H, et al. Spatial Clustering of de Novo Missense Mutations Identifies Candidate Neurodevelopmental Disorder-Associated Genes. *Am J Hum Genet.* 2017;101(3):478-484. doi:10.1016/j.ajhg.2017.08.004
28. Ferry Q, Steinberg J, Webber C, et al. Diagnostically relevant facial gestalt information from ordinary photos. *Elife.* 2014;3:e02020.
29. Reijnders MRF, Janowski R, Alvi M, et al. PURA syndrome: clinical delineation and genotype-phenotype study in 32 individuals with review of published literature. *J Med Genet.* 2018;55(2):104-113. doi:10.1136/jmedgenet-2017-104946
30. Murr R, Vaissière T, Sawan C, Shukla V, Herceg Z. Orchestration of chromatin-based processes: mind the TRRAP. *Oncogene.* 2007;26(37):5358-5372. doi:10.1038/sj.onc.1210605
31. Chen PB, Hung J-H, Hickman TL, et al. Hdac6 regulates Tip60-p400 function in stem cells. *Elife.* 2013;2:e01557. doi:10.7554/eLife.01557
32. Fazio TG, Huff JT, Panning B. An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. *Cell.* 2008;134(1):162-174. doi:10.1016/j.cell.2008.05.031
33. Acharya D, Hainer SJ, Yoon Y, et al. KAT-Independent Gene Regulation by Tip60 Promotes ESC Self-Renewal but Not Pluripotency. *Cell Rep.* 2017;19(4):671-679. doi:10.1016/j.celrep.2017.04.001
34. Tapias A, Wang Z-Q. Lysine Acetylation and Deacetylation in Brain Development and Neuropathies. *Genomics Proteomics Bioinformatics.* 2017;15(1):19-36. doi:10.1016/j.gpb.2016.09.002
35. Clayton-Smith J, O'Sullivan J, Daly S, et al. Whole-exome-sequencing identifies mutations in histone acetyltransferase gene KAT6B in individuals with the Say-Barber-Biesecker variant of Ohdo syndrome. *Am J Hum Genet.* 2011;89(5):675-681. doi:10.1016/j.ajhg.2011.10.008
36. Szakszon K, Salpietro C, Kakar N, et al. De novo mutations of the gene encoding the histone acetyltransferase KAT6B in two patients with Say-Barber/Biesecker/Young-Simpson syndrome. *Am J Med Genet A.* 2013;161A(4):884-888. doi:10.1002/ajmg.a.35848
37. Yilmaz R, Beleza-Meireles A, Price S, et al. A recurrent synonymous KAT6B mutation causes Say-Barber-Biesecker/Young-Simpson syndrome by inducing aberrant splicing. *Am J Med Genet A.* 2015;167A(12):3006-3010. doi:10.1002/ajmg.a.37343
38. Campeau PM, Kim JC, Lu JT, et al. Mutations in KAT6B, encoding a histone acetyltransferase, cause Genitopatellar syndrome. *Am J Hum Genet.* 2012;90(2):282-289. doi:10.1016/j.ajhg.2011.11.023
39. Simpson MA, Deshpande C, Dafou D, et al. De novo mutations of the gene encoding the histone acetyltransferase KAT6B cause Genitopatellar syndrome. *Am J Hum Genet.* 2012;90(2):290-294. doi:10.1016/j.ajhg.2011.11.024

40. Millan F, Cho MT, Retterer K, et al. Whole exome sequencing reveals de novo pathogenic variants in KAT6A as a cause of a neurodevelopmental disorder. *Am J Med Genet A*. 2016;170(7):1791-1798. doi:10.1002/ajmg.a.37670
41. Murray CR, Abel SN, McClure MB, et al. Novel Causative Variants in DYRK1A, KARS, and KAT6A Associated with Intellectual Disability and Additional Phenotypic Features. *J Pediatr Genet*. 2017;6(2):77-83. doi:10.1055/s-0037-1598639
42. Yan K, Rousseau J, Littlejohn RO, et al. Mutations in the Chromatin Regulator Gene BRPF1 Cause Syndromic Intellectual Disability and Deficient Histone Acetylation. *Am J Hum Genet*. 2017;100(1):91-104. doi:10.1016/j.ajhg.2016.11.011
43. Roelfsema JH, White SJ, Ariyürek Y, et al. Genetic heterogeneity in Rubinstein-Taybi syndrome: mutations in both the CBP and EP300 genes cause disease. *Am J Hum Genet*. 2005;76(4):572-580. doi:10.1086/429130
44. Bartholdi D, Roelfsema JH, Papadia F, et al. Genetic heterogeneity in Rubinstein-Taybi syndrome: delineation of the phenotype of the first patients carrying mutations in EP300. *J Med Genet*. 2007;44(5):327-333. doi:10.1136/jmg.2006.046698
45. Tsai AC-H, Dossett CJ, Walton CS, et al. Exon deletions of the EP300 and CREBBP genes in two children with Rubinstein-Taybi syndrome detected by aCGH. *Eur J Hum Genet*. 2011;19(1):43-49. doi:10.1038/ejhg.2010.121
46. Negri G, Milani D, Colapietro P, et al. Clinical and molecular characterization of Rubinstein-Taybi syndrome patients carrying distinct novel mutations of the EP300 gene. *Clin Genet*. 2015;87(2):148-154. doi:10.1111/cge.12348
47. Xu L-M, Li J-R, Huang Y, Zhao M, Tang X, Wei L. AutismKB: an evidence-based knowledgebase of autism genetics. *Nucleic Acids Res*. 2012;40(Database issue):D1016-1022. doi:10.1093/nar/gkr1145
48. Piton A, Jouan L, Rochefort D, et al. Analysis of the effects of rare variants on splicing identifies alterations in GABAA receptor genes in autism spectrum disorder individuals. *Eur J Hum Genet*. 2013;21(7):749-756. doi:10.1038/ejhg.2012.243
49. Lang B, Alrahbeni TMA, Clair DS, et al. HDAC9 is implicated in schizophrenia and expressed specifically in post-mitotic neurons but not in adult neural stem cells. *Am J Stem Cells*. 2012;1(1):31-41.
50. Severinsen JE, Bjarkam CR, Kiaer-Larsen S, et al. Evidence implicating BRD1 with brain development and susceptibility to both schizophrenia and bipolar affective disorder. *Mol Psychiatry*. 2006;11(12):1126-1138. doi:10.1038/sj.mp.4001885
51. Gause M, Eissenberg JC, MacRae AF, Dorsett M, Misulovin Z, Dorsett D. Nipped-A, the Tra1/TRRAP Subunit of the Drosophila SAGA and Tip60 Complexes, Has Multiple Roles in Notch Signaling during Wing Development. *Mol Cell Biol*. 2006;26(6):2347-2359. doi:10.1128/MCB.26.6.2347-2359.2006
52. Ceol CJ, Horvitz HR. A new class of C. elegans synMuv genes implicates a Tip60/NuA4-like HAT complex as a negative regulator of Ras signaling. *Dev Cell*. 2004;6(4):563-576.

53. Finkbeiner MG, Sawan C, Ouzounova M, Murr R, Herceg Z. HAT cofactor TRRAP mediates beta-catenine ubiquitination on the chromatin and the regulation of the canonical Wnt pathway. *Cell Cycle*. 2008;7(24):3908-3914. doi:10.4161/cc.7.24.7354
 54. Knutson BA, Hahn S. Domains of Tra1 important for activator recruitment and transcription coactivator functions of SAGA and NuA4 complexes. *Mol Cell Biol*. 2011;31(4):818-831. doi:10.1128/MCB.00687-10
 55. Driskell RR, Watt FM. Understanding fibroblast heterogeneity in the skin. *Trends Cell Biol*. 2015;25(2):92-99. doi:10.1016/j.tcb.2014.10.001
-
2. Venkatesh, S., and Workman, J.L. (2015). Histone exchange, chromatin structure and the regulation of transcription. *Nat. Rev. Mol. Cell Biol.* 16, 178–189.
 3. Hunt, C.R., Ramnarain, D., Horikoshi, N., Iyengar, P., Pandita, R.K., Shay, J.W., and Pandita, T.K. (2013). Histone Modifications and DNA Double-Strand Break Repair after Exposure to Ionizing Radiations. *Radiat. Res.* 179, 383–392.
 4. Legube, G., and Trouche, D. (2003). Regulating histone acetyltransferases and deacetylases. *EMBO Rep.* 4, 944–947.
 5. Berndsen, C.E., and Denu, J.M. (2008). Catalysis and substrate selection by histone/protein lysine acetyltransferases. *Curr. Opin. Struct. Biol.* 18, 682–689.
 6. McMahon, S.B., Van Buskirk, H.A., Dugan, K.A., Copeland, T.D., and Cole, M.D. (1998). The novel ATM-related protein TRRAP is an essential cofactor for the c-Myc and E2F oncoproteins. *Cell* 94, 363–374.
 7. Vassilev, A., Yamauchi, J., Kotani, T., Prives, C., Avantaggiati, M.L., Qin, J., and Nakatani, Y. (1998). The 400 kDa subunit of the PCAF histone acetylase complex belongs to the ATM superfamily. *Mol. Cell* 2, 869–875.
 8. Park, J., Kunjibettu, S., McMahon, S.B., and Cole, M.D. (2001). The ATM-related domain of TRRAP is required for histone acetyltransferase recruitment and Myc-dependent oncogenesis. *Genes Dev.* 15, 1619–1624.
 9. Zhao, L.-J., Loewenstein, P.M., and Green, M. (2017). Enhanced MYC association with the NuA4 histone acetyltransferase complex mediated by the adenovirus E1A N-terminal domain activates a subset of MYC target genes highly expressed in cancer cells. *Genes Cancer* 8, 752–761.
 10. Jethwa, A., Ślabicki, M., Hüllelin, J., Jentzsch, M., Dalal, V., Rabe, S., Wagner, L., Walther, T., Klapper, W., Bohnenberger, H., et al. (2018). TRRAP is essential for regulating the accumulation of mutant and wild-type p53 in lymphoma. *Blood*.
 11. Wei, X., Walia, V., Lin, J.C., Teer, J.K., Prickett, T.D., Gartner, J., Davis, S., NISC Comparative Sequencing Program, Stemke-Hale, K., Davies, M.A., et al. (2011). Exome sequencing identifies GRIN2A as frequently mutated in melanoma. *Nat. Genet.* 43, 442–446.

12. Wurdak, H., Zhu, S., Romero, A., Lorger, M., Watson, J., Chiang, C.-Y., Zhang, J., Natu, V.S., Lairson, L.L., Walker, J.R., et al. (2010). An RNAi screen identifies TRRAP as a regulator of brain tumor-initiating cell differentiation. *Cell Stem Cell* 6, 37–47.
13. Loukopoulos, P., Shibata, T., Katoh, H., Kokubu, A., Sakamoto, M., Yamazaki, K., Kosuge, T., Kanai, Y., Hosoda, F., Imoto, I., et al. (2007). Genome-wide array-based comparative genomic hybridization analysis of pancreatic adenocarcinoma: identification of genetic indicators that predict patient outcome. *Cancer Sci.* 98, 392–400.
14. Herceg, Z., Hulla, W., Gell, D., Cuenin, C., Lleonart, M., Jackson, S., and Wang, Z.Q. (2001). Disruption of *Trrap* causes early embryonic lethality and defects in cell cycle progression. *Nat. Genet.* 29, 206–211.
15. Sawan, C., Hernandez-Vargas, H., Murr, R., Lopez, F., Vaissière, T., Ghantous, A.Y., Cuenin, C., Imbert, J., Wang, Z.-Q., Ren, B., et al. (2013). Histone acetyltransferase cofactor *Trrap* maintains self-renewal and restricts differentiation of embryonic stem cells. *Stem Cells Dayt. Ohio* 31, 979–991.
16. Tapias, A., Zhou, Z.-W., Shi, Y., Chong, Z., Wang, P., Groth, M., Platzer, M., Huttner, W., Herceg, Z., Yang, Y.-G., et al. (2014). *Trrap*-dependent histone acetylation specifically regulates cell-cycle gene transcription to control neural progenitor fate decisions. *Cell Stem Cell* 14, 632–643.
17. Gupta, A.R., Westphal, A., Yang, D.Y.J., Sullivan, C.A.W., Eilbott, J., Zaidi, S., Voos, A., Vander Wyk, B.C., Ventola, P., Waqar, Z., et al. (2017). Neurogenetic analysis of childhood disintegrative disorder. *Mol. Autism* 8, 19.
18. Xu, B., Ionita-Laza, I., Roos, J.L., Boone, B., Woodrick, S., Sun, Y., Levy, S., Gogos, J.A., and Karayiorgou, M. (2012). De novo gene mutations highlight patterns of genetic and neural complexity in schizophrenia. *Nat. Genet.* 44, 1365–1369.
19. Takata, A., Xu, B., Ionita-Laza, I., Roos, J.L., Gogos, J.A., and Karayiorgou, M. (2014). Loss-of-Function Variants in Schizophrenia Risk and *SETD1A* as a Candidate Susceptibility Gene. *Neuron* 82, 773–780.
20. Geisheker, M.R., Heymann, G., Wang, T., Coe, B.P., Turner, T.N., Stessman, H.A.F., Hoekzema, K., Kvarnung, M., Shaw, M., Friend, K., et al. (2017). Hotspots of missense mutation identify neurodevelopmental disorder genes and functional domains. *Nat. Neurosci.* 20, 1043–1051.
21. Sobreira, N., Schiettecatte, F., Valle, D., and Hamosh, A. (2015). GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. *Hum. Mutat.* 36, 928–930.
22. Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536, 285–291.
23. Kircher, M., Witten, D.M., Jain, P., O'Roak, B.J., Cooper, G.M., and Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* 46, 310–315.

24. Kumar, P., Henikoff, S., and Ng, P.C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* 4, 1073–1081.
25. Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., and Sunyaev, S.R. (2010). A method and server for predicting damaging missense mutations. *Nat. Methods* 7, 248–249.
26. Ware, J.S., Samocha, K.E., Homsy, J., and Daly, M.J. (2015). Interpreting de novo variation in human disease using denovolyzeR. *Curr. Protoc. Hum. Genet.* Editor. Board Jonathan Haines A1 87, 7.25.1-7.25.15.
27. Lelieveld, S.H., Wiel, L., Venselaar, H., Pfundt, R., Vriend, G., Veltman, J.A., Brunner, H.G., Vissers, L.E.L.M., and Gilissen, C. (2017). Spatial Clustering of de Novo Missense Mutations Identifies Candidate Neurodevelopmental Disorder-Associated Genes. *Am. J. Hum. Genet.* 101, 478–484.
28. Ferry, Q., Steinberg, J., Webber, C., FitzPatrick, D.R., Ponting, C.P., Zisserman, A., and Nellåker, C. (2014). Diagnostically relevant facial gestalt information from ordinary photos. *ELife* 3, e02020.
29. Reijnders, M.R.F., Janowski, R., Alvi, M., Self, J.E., van Essen, T.J., Vreeburg, M., Rouhl, R.P.W., Stevens, S.J.C., Stegmann, A.P.A., Schieving, J., et al. (2018). PURA syndrome: clinical delineation and genotype-phenotype study in 32 individuals with review of published literature. *J. Med. Genet.* 55, 104–113.
30. Murr, R., Vaissière, T., Sawan, C., Shukla, V., and Herceg, Z. (2007). Orchestration of chromatin-based processes: mind the TRRAP. *Oncogene* 26, 5358–5372.
31. Chen, P.B., Hung, J.-H., Hickman, T.L., Coles, A.H., Carey, J.F., Weng, Z., Chu, F., and Fazio, T.G. (2013). Hdac6 regulates Tip60-p400 function in stem cells. *ELife* 2, e01557.
32. Fazio, T.G., Huff, J.T., and Panning, B. (2008). An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. *Cell* 134, 162–174.
33. Acharya, D., Hainer, S.J., Yoon, Y., Wang, F., Bach, I., Rivera-Pérez, J.A., and Fazio, T.G. (2017). KAT-Independent Gene Regulation by Tip60 Promotes ESC Self-Renewal but Not Pluripotency. *Cell Rep.* 19, 671–679.
34. Tapias, A., and Wang, Z.-Q. (2017). Lysine Acetylation and Deacetylation in Brain Development and Neuropathies. *Genomics Proteomics Bioinformatics* 15, 19–36.
35. Clayton-Smith, J., O’Sullivan, J., Daly, S., Bhaskar, S., Day, R., Anderson, B., Voss, A.K., Thomas, T., Biesecker, L.G., Smith, P., et al. (2011). Whole-exome-sequencing identifies mutations in histone acetyltransferase gene KAT6B in individuals with the Say-Barber-Biesecker variant of Ohdo syndrome. *Am. J. Hum. Genet.* 89, 675–681.
36. Szakszon, K., Salpietro, C., Kakar, N., Knecht, A.C., Oláh, É., Dallapiccola, B., and Borck, G. (2013). De novo mutations of the gene encoding the histone acetyltransferase KAT6B in two patients with Say-Barber/Biesecker/Young-Simpson syndrome. *Am. J. Med. Genet. A.* 161A, 884–888.

37. Yilmaz, R., Beleza-Meireles, A., Price, S., Oliveira, R., Kubisch, C., Clayton-Smith, J., Szakszon, K., and Borck, G. (2015). A recurrent synonymous KAT6B mutation causes Say-Barber-Biesecker/Young-Simpson syndrome by inducing aberrant splicing. *Am. J. Med. Genet. A.* *167A*, 3006–3010.
38. Campeau, P.M., Kim, J.C., Lu, J.T., Schwartzentruber, J.A., Abdul-Rahman, O.A., Schlaubitz, S., Murdock, D.M., Jiang, M.-M., Lammer, E.J., Enns, G.M., et al. (2012). Mutations in KAT6B, encoding a histone acetyltransferase, cause Genitopatellar syndrome. *Am. J. Hum. Genet.* *90*, 282–289.
39. Simpson, M.A., Deshpande, C., Dafou, D., Vissers, L.E.L.M., Woollard, W.J., Holder, S.E., Gillessen-Kaesbach, G., Derks, R., White, S.M., Cohen-Snuijf, R., et al. (2012). De novo mutations of the gene encoding the histone acetyltransferase KAT6B cause Genitopatellar syndrome. *Am. J. Hum. Genet.* *90*, 290–294.
40. Millan, F., Cho, M.T., Retterer, K., Monaghan, K.G., Bai, R., Vitazka, P., Everman, D.B., Smith, B., Angle, B., Roberts, V., et al. (2016). Whole exome sequencing reveals de novo pathogenic variants in KAT6A as a cause of a neurodevelopmental disorder. *Am. J. Med. Genet. A.* *170*, 1791–1798.
41. Murray, C.R., Abel, S.N., McClure, M.B., Foster, J., Walke, M.I., Jayakar, P., Bademci, G., and Tekin, M. (2017). Novel Causative Variants in DYRK1A, KARS, and KAT6A Associated with Intellectual Disability and Additional Phenotypic Features. *J. Pediatr. Genet.* *6*, 77–83.
42. Yan, K., Rousseau, J., Littlejohn, R.O., Kiss, C., Lehman, A., Rosenfeld, J.A., Stumpel, C.T.R., Stegmann, A.P.A., Robak, L., Scaglia, F., et al. (2017). Mutations in the Chromatin Regulator Gene BRPF1 Cause Syndromic Intellectual Disability and Deficient Histone Acetylation. *Am. J. Hum. Genet.* *100*, 91–104.
43. Roelfsema, J.H., White, S.J., Ariyürek, Y., Bartholdi, D., Niedrist, D., Papadia, F., Bacino, C.A., den Dunnen, J.T., van Ommen, G.-J.B., Breuning, M.H., et al. (2005). Genetic heterogeneity in Rubinstein-Taybi syndrome: mutations in both the CBP and EP300 genes cause disease. *Am. J. Hum. Genet.* *76*, 572–580.
44. Bartholdi, D., Roelfsema, J.H., Papadia, F., Breuning, M.H., Niedrist, D., Hennekam, R.C., Schinzel, A., and Peters, D.J.M. (2007). Genetic heterogeneity in Rubinstein-Taybi syndrome: delineation of the phenotype of the first patients carrying mutations in EP300. *J. Med. Genet.* *44*, 327–333.
45. Tsai, A.C.-H., Dossett, C.J., Walton, C.S., Cramer, A.E., Eng, P.A., Nowakowska, B.A., Pursley, A.N., Stankiewicz, P., Wiszniewska, J., and Cheung, S.W. (2011). Exon deletions of the EP300 and CREBBP genes in two children with Rubinstein-Taybi syndrome detected by aCGH. *Eur. J. Hum. Genet. EJHG* *19*, 43–49.
46. Negri, G., Milani, D., Colapietro, P., Forzano, F., Della Monica, M., Rusconi, D., Consonni, L., Caffi, L.G., Finelli, P., Scarano, G., et al. (2015). Clinical and molecular characterization of Rubinstein-Taybi syndrome patients carrying distinct novel mutations of the EP300 gene. *Clin. Genet.* *87*, 148–154.

47. Xu, L.-M., Li, J.-R., Huang, Y., Zhao, M., Tang, X., and Wei, L. (2012). AutismKB: an evidence-based knowledgebase of autism genetics. *Nucleic Acids Res.* *40*, D1016-1022.
48. Piton, A., Jouan, L., Rochefort, D., Dobrzeniecka, S., Lachapelle, K., Dion, P.A., Gauthier, J., and Rouleau, G.A. (2013). Analysis of the effects of rare variants on splicing identifies alterations in GABAA receptor genes in autism spectrum disorder individuals. *Eur. J. Hum. Genet.* *21*, 749–756.
49. Lang, B., Alrahbeni, T.M.A., Clair, D.S., Blackwood, D.H., International Schizophrenia Consortium, McCaig, C.D., and Shen, S. (2012). HDAC9 is implicated in schizophrenia and expressed specifically in post-mitotic neurons but not in adult neural stem cells. *Am. J. Stem Cells* *1*, 31–41.
50. Severinsen, J.E., Bjarkam, C.R., Kiaer-Larsen, S., Olsen, I.M., Nielsen, M.M., Blechingberg, J., Nielsen, A.L., Holm, I.E., Foldager, L., Young, B.D., et al. (2006). Evidence implicating BRD1 with brain development and susceptibility to both schizophrenia and bipolar affective disorder. *Mol. Psychiatry* *11*, 1126–1138.
51. Gause, M., Eissenberg, J.C., MacRae, A.F., Dorsett, M., Misulovin, Z., and Dorsett, D. (2006). Nipped-A, the Tra1/TRRAP Subunit of the Drosophila SAGA and Tip60 Complexes, Has Multiple Roles in Notch Signaling during Wing Development. *Mol. Cell. Biol.* *26*, 2347–2359.
52. Ceol, C.J., and Horvitz, H.R. (2004). A new class of *C. elegans* synMuv genes implicates a Tip60/NuA4-like HAT complex as a negative regulator of Ras signaling. *Dev. Cell* *6*, 563–576.
53. Finkbeiner, M.G., Sawan, C., Ouzounova, M., Murr, R., and Herceg, Z. (2008). HAT cofactor TRRAP mediates beta-catenine ubiquitination on the chromatin and the regulation of the canonical Wnt pathway. *Cell Cycle Georget. Tex* *7*, 3908–3914.
54. Knutson, B.A., and Hahn, S. (2011). Domains of Tra1 important for activator recruitment and transcription coactivator functions of SAGA and NuA4 complexes. *Mol. Cell. Biol.* *31*, 818–831.
55. Zhou, X., Edmonson, M.N., Wilkinson, M.R., Patel, A., Wu, G., Liu, Y., Li, Y., Zhang, Z., Rusch, M.C., Parker, M., et al. (2016). Exploring genomic alteration in pediatric cancer using ProteinPaint. *Nat. Genet.* *48*, 4–6.
56. Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* *10*, 845–858.
57. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* *25*, 1605–1612.

Supplementary Material

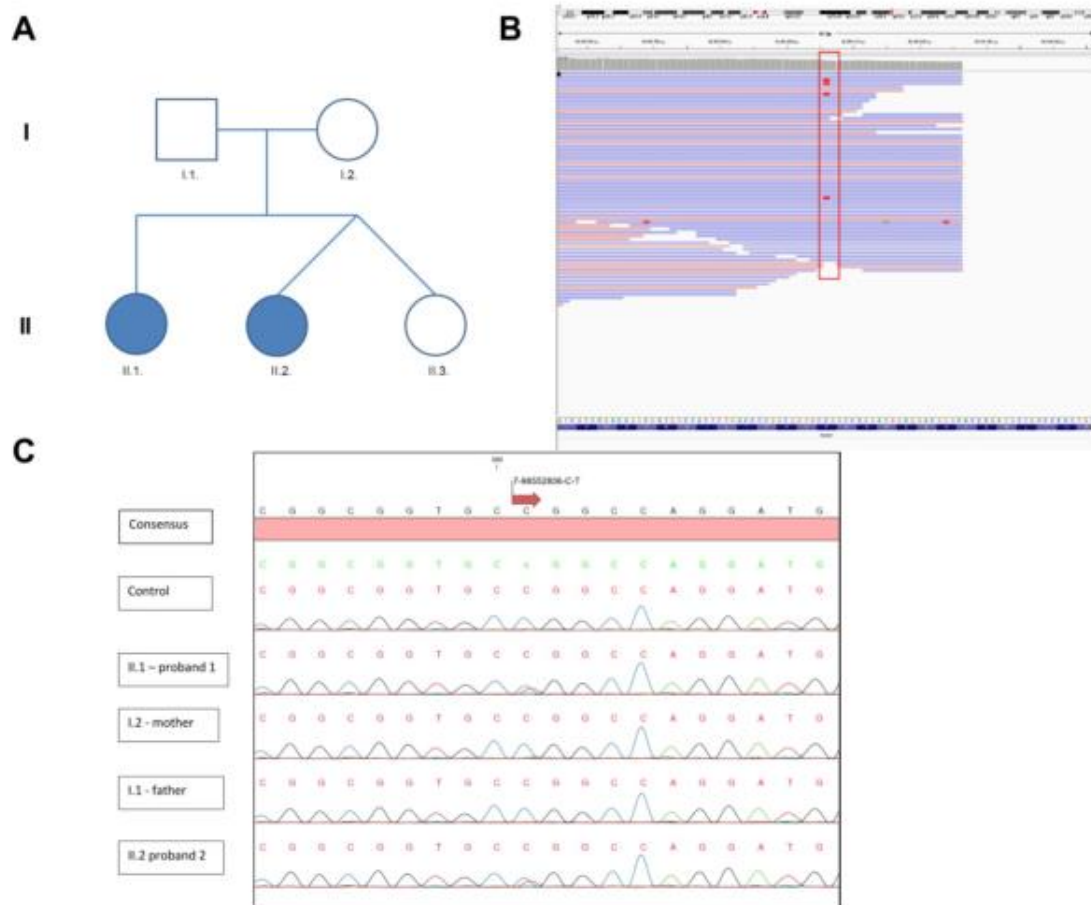


Figure S1. Maternal mosaicism of variant c.5795C>T (p.Pro1932Leu) identified in individuals 22 and 23.

(A) Pedigree of the family. (B) Integrative Genome Viewer snapshot of BAM emitted by GATKhaplotype caller in the mother (individual I.2.). Reads are shown aligned to the genome in blue for forwardstrand and red for reverse strand. The C to T variant is shown in red on 4/55 reads. This low 7% representation is significantly deviated from the expected 50% and is indicative of mosaicism. (C) Segregation analysis of the variant by Sanger sequencing confirms the heterozygous variant in the siblings and the mosaicism in the mother.

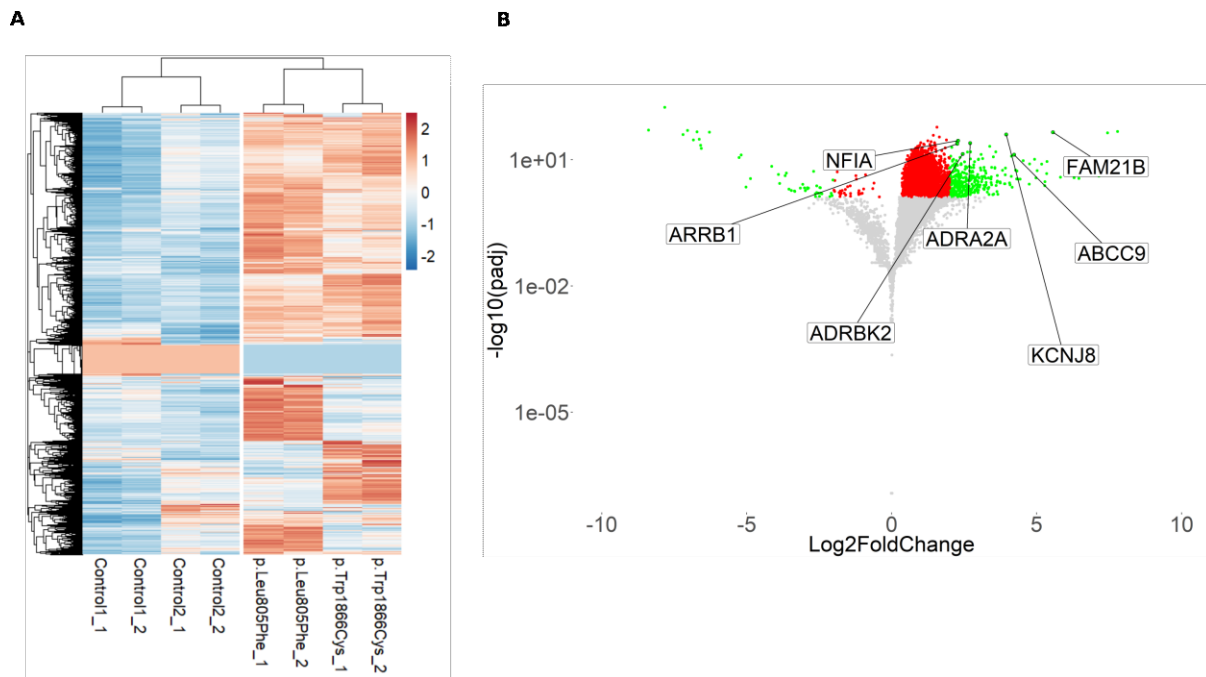


Figure S2. Fibroblasts harboring the p.(Leu805Phe) or p.(Trp186Cys) variants have differential gene expression patterns.

RNAseq was performed on two healthy controls and two individual fibroblasts to assay gene expression with technical duplicates. (A) Heatmap of genes expressed with at least 10 counts in at least one condition. Normalized count values from “DESeq2” were plotted and scaled row-wise using the “pheatmap” R package. (B) Log2Fold Change (Log2FC) and p-values adjusted for 10% False Discovery Rate (padj) were calculated for the two affected individuals and two controls respectively pooled together using the “DESeq2” R package then plotted using the “ggplot2” R package as a volcano plot. Green dots represent genes with a padj lower than 0.01, and an absolute Log2FC higher than 2 ($\text{padj} < 0.01$, $\text{abs}(\text{Log2FC}) > 2$) and are referred to as Differentially Expressed Genes (DEGs). Red dots represent genes with a padj lower than 0.01 but an absolute Log2FC lower than 2 ($\text{padj} < 0.01$, $\text{abs}(\text{Log2FC}) < 2$).

Supplementary methods

RNAseq

Human primary fibroblasts for patient were cultured in DMEM (ThermoFisher cat# 11995-065), 10% FBS, 1mM GlutaMax (ThermoFisher cat# 35050-061) and antibiotics-antimycotics (ThermoFisher cat# 15240-062). Fibroblasts were plated at 1 million

cells per 150mm dish and allowed to grow until they reached 80% confluency. Cells were washed twice with D-PBS, resuspended in QIAzol (Qiagen cat# 79306) and stored at -80°C until all samples were ready for RNA extraction. RNA isolation was performed using the RNeasy mini kit (Qiagen cat# 74104), according to the manufacturer's protocol. Samples were treated with the Turbo DNA free kit (ThermoFisher cat# AM1907) and quality was assessed using the Agilent 2100 Bioanalyzer. Sequencing was performed at the CHU Sainte-Justine and Génome Québec Integrated Clinical Genomic Centre in Pediatrics (CIGCP). mRNA Libraries were prepared using the TruSeq Stranded mRNA kit for 48 samples (Illumina), according to the manufacturer's instructions. Samples were run on the Illumina HiSeq 4000 PE100 with 7 samples per lane. Output files were analyzed using the MUGQIC RNAseq pipeline (MUGQIC) steps 1 through 14 on the Guillimin Génome Québec HPC. In summary, BAM files were converted to FASTQ using Picard (BROAD Institute), sequences were trimmed using Trimmomatic (Bolger, 2014) then aligned to the GRCh37 genome using STAR (Dobin, 2013); duplicate and misaligned reads were discarded using Picard, and raw counts were called using HTseq (Anders S, 2015). Differential expression analysis was performed using the DESeq2 R package (Love, 2014), with default parameters. GO annotation analysis was performed using the GOrilla web application (Eden, 2009). Control and patient cell lines of the same type were respectively pooled together into two groups for differential gene expression analysis, so as to find genes which were significantly differentially expressed in all patients compared to all controls. Significantly expressed genes were selected with an adjusted p value (10% False Discovery Rate, padj) lower than 0.01, and a log2 Fold Change (log2FC) higher than 2 or lower than -2; corresponding to an overall fold change of at least 4 or -4. Genes of interest were all significantly differentially expressed in individual patient analyses with DESeq2 compared to controls.

Bibliography for methods section

1. Anders S, P. P. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics*, 31(2), 166-9.
2. Bolger, A. M. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120.

3. BROAD Institute. (n.d.). Picard Tools. Retrieved 2017, from <https://github.com/broadinstitute/picard>
4. Dobin, A. D. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1), 15–21.
5. Eden, E. N. (2009). GOrilla: A Tool For Discovery And Visualization of Enriched GO Terms in Ranked Gene Lists. *BMC Bioinformatics*, 10(48).
6. Love, M. I. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12), 550.
7. MUGQIC. (n.d.). MUGQIC pipelines RNAseq, 2.2.1-beta. Retrieved 2017, from https://bitbucket.org/mugqic/mugqic_pipelines/src/master/pipelines/rnaseq/

3.1.2. Vérification des résultats de RNAseq

Afin de vérifier les résultats de séquençage, on a effectué des analyses ciblées sur les gènes les plus différentiellement régulés, et qui auraient un intérêt pertinent par rapport au phénotype neurodéveloppemental des individus affectés. Pour obtenir des résultats plus robustes, nous avons utilisé des contrôles différents de ceux utilisés en RNAseq, et en plus grande nombre. Aussi, comme nous continuons le recrutement d'individus affectés, nous avons reçu plus de cellules après le séquençage sur lesquelles nous avons pu tester les gènes sélectionnés pour les deux premières lignées (Figure 10). Sachant que l'un des termes GO (Gene Ontology) qui était le plus enrichi dans les gènes différentiellement régulés était les transporteurs d'ions, nous nous sommes concentrés sur ceux-ci. Malheureusement, les fibroblastes de l'individu porteur de la mutation W1866C étaient devenus sénescents, mais nous avons pu tester l'individu porteur de la mutation L805F, et nous avons rajouté la mutation G1883R et un individu du regroupement plus syndromique porteur de la mutation récurrente A1043T. Sachant que les individus testés étaient de sexe différent (2 femelles, 1 mâle), d'âge différent (2, 8, et 11 ans), et que les passages des cellules étaient variés, nous avons utilisé des contrôles des deux sexes et de passages les plus proches des cellules utilisées (entre P5 et P10 à partir de l'établissement des cellules). Alors, nous nous sommes heurtés au même problème que pour le RNAseq, c'est-à-dire la variabilité des contrôles qui, même en augmentant à cinq lignées différentes, présentent des écart-types relativement hauts. Pour deux des gènes dont la *padj* était la plus significative, soit *ABCC9* et *KCNJ8*, on observe une tendance non significative pour *ABCC9* et une augmentation significative pour *KCNJ8* vers une

surexpression de ces gènes dans les individus affectés qui ne font pas partie du regroupement plus syndromique (L805F, G1883R) (Figure 10). Le manque de reproductibilité pour les autres gènes semblerait provenir d'un manque de contrôles lors de l'expérience de RNAseq, qui induirait un biais et donnerait lieu à des faux positifs. Cependant, le manque de reproductibilité du phénomène avec les cellules de l'individu porteur d'une mutation dans le regroupement plus syndromique pourrait laisser penser que différents gènes sont affectés dû aux positions différentes des mutations qui affecteraient des interactions différentes de la protéine TRRAP, et donc des voies biologiques différentes.

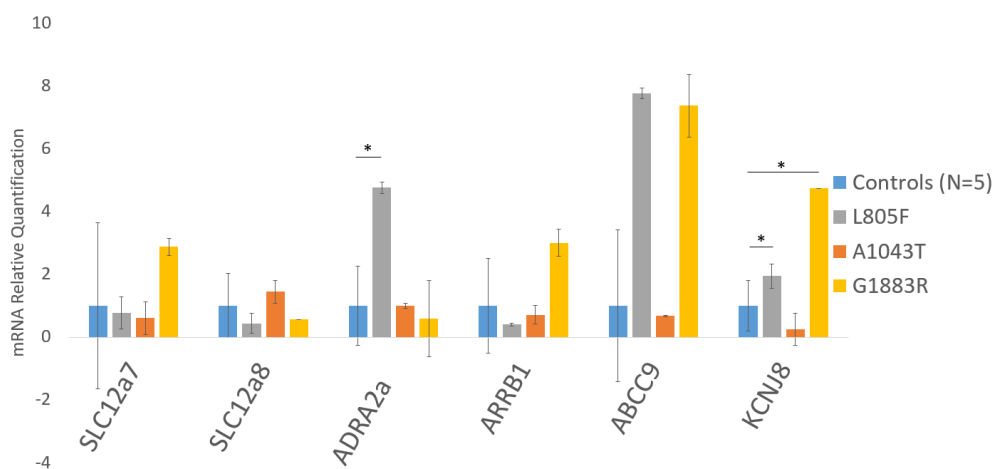


Figure 10: Vérification des gènes sélectionnés du RNAseq par qPCR. L'ARNm a été extrait de cinq fibroblastes sains et des individus affectés en quadruplicatas. L'expression de chaque gène est normalisée à l'expression de *GAPDH*. * : $p < 0.05$ avec un test T pour échantillons indépendants.

3.1.3 Séquençage de l'ARNm de LCLs d'individus TRRAP

En plus des fibroblastes d'individus TRRAP, nous avons aussi obtenu des lignées lymphoblastoïques pour trois individus : l'individu porteur de la mutation L805F, et deux individus porteurs de la mutations récurrente A1043T. En les comparant à trois lignées contrôles, on observe qu'il existe 64 gènes différentiellement régulés dans les cellules d'individus TRRAP (Figure 11). Sur ce total, seulement cinq gènes sont aussi retrouvés dans les DEGs des fibroblastes TRRAP (Annexe 1), et de ceux-ci un seul est différentiellement régulé dans le même sens. Cependant, une grande partie des DEGs des LCLs est impliquée dans le fonctionnement neuronal, le développement neural, ou le cytosquelette ou l'adhésion cellulaire (Tableau V). Alors, il est possible que TRRAP interagisse avec des facteurs de

transcription ou des sous-unités de complexes de remodelage de la chromatine qui soient spécifiques au type-cellulaire, et qui auraient des *loci* cibles différents.

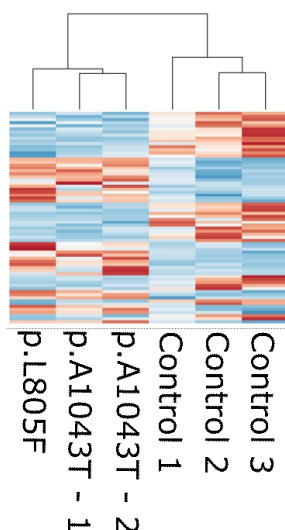


Figure 11 : Les LCLs d'individus TRRAP présentent des gènes différentiellement régulés. L'ARN de LCLs provenant de 3 individus affectés et 3 contrôles sains a été extrait et des libraires pour séquençage Illumina ont été préparées à partir de 800ng d'ARN. Heatmap des DEGs avec $p < 0.01$ et $\log_2FC > 2$ entre les individus TRRAP et les contrôles sains. L'expression des gènes plus élevée est représentée par la couleur rouge, tandis que l'expression moins élevée est représentée par la couleur bleue.

Tableau V - Gènes différentiellement régulés dans les LCLs d'individus porteurs de variants dans TRRAP impliqués dans le développement neural. Les gènes significativement dérégulés dans les 3 LCLs d'individus affectés porteurs de variants dans TRRAP ont été catégorisés par annotation manuelle et seuls ceux ayant un rôle dans la fonction neuronale, le développement neural, ou l'adhérence cellulaire ou le cytosquelette ont été sélectionnés.

| Gène | log2FoldChange | padj | Fonction |
|------------|----------------|----------|------------------------------------|
| CACNA1E | -3.02687 | 3.49E-08 | Fonction Neuronale |
| CHI3L1 | -4.57399 | 1.96E-05 | Cytosquelette/Adhérence Cellulaire |
| CSGALNACT1 | 3.950611 | 3.00E-05 | Cytosquelette/Adhérence Cellulaire |
| TNC | -3.02106 | 3.55E-05 | Développement Neural |
| CPXM2 | -4.22239 | 3.64E-05 | Fonction Neuronale |
| FAT2 | -3.31768 | 3.91E-05 | Cytosquelette/Adhérence Cellulaire |
| ZNF730 | 2.593326 | 6.03E-05 | Fonction Neuronale |
| SCHIP1 | 3.867895 | 0.001043 | Développement Neural |
| RAB6B | -4.44294 | 0.001227 | Fonction Neuronale |
| RASGRF1 | -2.6115 | 0.001679 | Fonction Neuronale |
| ADORA2B | 2.766366 | 0.002155 | Développement Neural |
| FHOD3 | -3.46566 | 0.002511 | Cytosquelette/Adhérence Cellulaire |

| | | | |
|-----------------|----------|----------|------------------------------------|
| IQCA1 | 3.374123 | 0.002511 | Cytosquelette/Adhérence Cellulaire |
| NTN4 | 4.149745 | 0.002511 | Développement Neural |
| NFIX | -2.05355 | 0.004128 | Développement Neural |
| FXYD6 | -4.29767 | 0.004134 | Fonction Neuronale |
| ARHGEF28 | -2.91451 | 0.004531 | Fonction Neuronale |
| CCDC85C | -2.94385 | 0.005947 | Développement Neural |
| PIGR | -2.74115 | 0.005998 | Cytosquelette/Adhérence Cellulaire |
| SOWAHB | -3.93411 | 0.006805 | Développement Neural |
| DHRS9 | 3.25298 | 0.007148 | Cytosquelette/Adhérence Cellulaire |
| FRMPD2 | -2.22181 | 0.007371 | Cytosquelette/Adhérence Cellulaire |
| KAL1 | 4.014948 | 0.007372 | Cytosquelette/Adhérence Cellulaire |
| FOXP2 | -3.23619 | 0.008543 | Développement Neural |
| SATB1 | 3.263034 | 0.00938 | Fonction Neuronale |
| HHIP | 3.30666 | 0.009555 | Développement Neural |
| MED12L | 2.877447 | 0.009555 | Fonction Neuronale |

3.2. Des mutations dans *SMARCC2* sont associées à un syndrome neurodéveloppemental et causent des changements d'expression génique

3.2.1 Mise en situation

Le gène *SMARCC2* encode SMARCC2, ou BAF170, une protéine structurale formant des homo- ou hétéro-dimères avec son homologue SMARCC1. Ce dimère et SMARCD1 forment le cœur du complexe BAF, ou SWI/SNF, sur lequel toutes les autres sous-unités se placent afin de former un complexe de remodelage des nucléosomes fonctionnel. Nous avons caractérisé des individus porteurs de mutations dans *SMARCC2* et présentant un syndrome neurodéveloppemental similaire aux autres syndromes associés à différentes sous unités du complexe BAF, le TDISS (Trouble de Déficience Intellectuelle associé au complexe SWI/SNF). Nous avons aussi démontré que certaines des mutations affectent l'expression génique dans des cellules de ces individus, et que plusieurs de ces gènes sont impliqués dans le développement ou le fonctionnement neural. Nous avons aussi démontré que certaines de ces mutations causaient des défauts d'épissage, soit l'exclusion d'un exon, ou la dégradation d'ARNm non-sens.

Keren Machol a procédé au recrutement et à la caractérisation clinique des individus, et a démontré l'exclusion exon de l'individu huit, et j'ai démontré la dégradation de l'ARNm

de *SMARCC2* dans l'individu quatre. L'établissement des lignées a été fait par Thi-Tuyet Mai Nguyen, Justine Rousseau, et moi-même. Les qPCR ont été réalisées par Thomas Garcia et la préparation pour le séquençage de l'ARNm a été fait par Justine Rousseau. J'ai effectué toutes les analyses bio-informatiques. Les autres auteurs sont des cliniciens qui ont participé au recrutement et à la collecte de données des individus affectés. Keren, Justine, et moi-même avons écrit le manuscrit publié : Machol et al., *Am J Hum Gen*, 2019¹⁹⁰.

3.2.2 Article 2

Expanding the Spectrum of BAF-related Disorders – *De Novo* Variants in *SMARCC2* Cause a Syndrome with Intellectual Disability and Developmental Delay

Keren Machol,¹ Justine Rousseau,² Sophie Ehresmann,² Thomas Garcia,² Thi Tuyet Mai Nguyen,² Rebecca C. Spillmann,³ Jennifer A. Sullivan,³ Vandana Shashi,³ Yong-Hui Jiang,⁴ Nicholas Stong,⁵ Elise Fiala,⁶ Marcia Willing,⁶ Rolph Pfundt,⁷ Tjitske Kleefstra,⁷ Megan T. Cho,⁸ Heather McLaughlin,⁸ Monica Rosello Piera,⁹ Carmen Orellana,⁹ Francisco Martínez,⁹ Alfonso Caro-Llopis,⁹ Sandra Monfort,⁹ Tony Roscioli,¹⁰ Cheng Yee Nixon,¹⁰ Michael Buckley,¹¹ Anne Turner,¹² Wendy Jones,¹³ Peter M. van Hasselt,¹⁴ Floris C. Hofstede,¹⁴ Koen L.I. van Gassen,¹⁴ Alice S. Brooks,¹⁵ Marjon A. van Slegtenhorst,¹⁵ Katherine Lachlan,¹⁶ Jessica Sebastian,¹⁷ Suneeta Madan-Khetarpal,¹⁷ Desai Sonal,¹⁸ Naidu Sakkubai,¹⁸ Thevenon Julien,¹⁹ Faivre Laurence,¹⁹ Masurel Alice,¹⁹ Slavé Petrovski,²⁰ , Ian D. Krantz,²¹ Jennifer M. Tarpinian,²¹ Jill A. Rosenfeld,¹ Brendan H. Lee,¹ Undiagnosed Diseases Network, Philippe M. Campeau^{2*}

1) Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, 77030, USA

- 2) Department of Pediatrics, CHU Sainte-Justine Research Center and University of Montreal, Montreal, QC, H3T 1C5, Canada
- 3) Department of Pediatrics, Duke University School of Medicine, Durham, North Carolina 27710, USA
- 4) Department of Pediatrics and Neurobiology, Program in Genetics and Genomics, Duke University School of Medicine, Durham, North Carolina 27710, USA
- 5) Institute for Genomic Medicine, Columbia University, New York, NY 10032, USA
- 6) Department of Pediatrics, Washington University School of Medicine, St. Louis, MO, USA
- 7) Human Genetics Department, Radboud University Medical Center, HB, Nijmegen, The Netherlands.
- 8) GeneDx, Gaithersburg, Maryland, USA
- 9) Unidad de Genética, Hospital Universitari i Politècnic La Fe, Valencia, Spain
- 10) Centre for Clinical Genetics, Sydney Children's Hospital, Neuroscience Research Australia, University of New South Wales, Sydney, Australia
- 11) New South Wales Health Pathology, Randwick, Australia
- 12) Centre for Clinical Genetics, Sydney Children's Hospital, Sydney, Australia
- 13) North East Thames Regional Genetics service, Great Ormond Street Hospital for Children NHS Foundation Trust, London WC1N 3JH, UK
- 14) Department of Metabolic Diseases, Wilhelmina Children's Hospital, University Medical Center Utrecht, The Netherlands
- 15) Department of Clinical Genetics, Erasmus Medical Center, Rotterdam, The Netherlands
- 16) Wessex Clinical Genetics Service, Princess Anne Hospital, Southampton, UK
- 17) Department of Medical Genetics, Children's Hospital of Pittsburgh of UPMC, Pittsburgh, PA

- 18) Department of Neurogenetics, Kennedy Krieger Institute, Baltimore, MD, USA
- 19) Centre de Génétique, Centre de Référence Anomalies du Développement et Syndromes Malformatifs de l'Interrégion Est et FHU TRANSLAD, CHU Dijon, F-21079 Dijon, France
- 20) AstraZeneca Centre for Genomics Research, Precision Medicine and Genomics, IMED Biotech Unit, AstraZeneca, Cambridge CB2 0AA, UK
- 21) Department of Pediatrics and Division of Human Genetics, The Children's Hospital of Philadelphia and the Perelman School of Medicine at The University of Pennsylvania, Philadelphia PA 19104, USA

* Correspondence: p.campeau@umontreal.ca

Abstract

SMARCC2 (BAF170, MIM: 601734) is one of the invariable core subunits of the ATP-dependent chromatin remodeling BAF (BRG1-Associated Factor) complex and plays a crucial role in embryogenesis and corticogenesis. Pathogenic variants in genes encoding other components of the BAF-complex have been associated with intellectual disability syndromes. Despite its significant biological role variants in *SMARCC2* have not been directly associated with human disease previously. Using whole exome sequencing and a web-based gene-matching program, we identified 15 individuals with variable degrees of neurodevelopmental delay and growth retardation harboring one of 13 heterozygous variants in *SMARCC2*, most of them novel and proven *de novo*. The clinical presentation overlaps with intellectual disability syndromes associated with other BAF subunits, such as Coffin Siris and Nicolaides–Baraitser syndromes, and includes prominent speech impairment, hypotonia, feeding difficulties, behavioral abnormalities, and dysmorphic features

such as hypertrichosis, thick eyebrows, thin upper lip vermilion, and upturned nose. Nine out of the fifteen individuals harbor variants in the highly conserved SMARCC2 DNA-interacting domains (SANT and SWIRM) and present with a more severe phenotype. Two of these individuals present cardiac abnormalities. Transcriptomic analysis of fibroblasts from affected individuals highlights a group of differentially expressed genes with possible roles in regulation of neuronal development and function, namely *H19*, *SCRG1*, *RELN* and *CACNB4*. Our findings suggest a novel SMARCC2-related syndrome which overlaps with neurodevelopmental disorders associated with variants in BAF-complex subunits.

Main Text

Introduction:

The chromatin-remodeling complex BRG1-Associated Factor (BAF) plays an essential role in the regulation of gene expression and higher-order chromatin organization by modulating the nucleosome and changing chromatin conformation and accessibility ^{1,2}. BAFopathies are a heterogeneous group of disorders caused by mutations in the various subunits composing the BAF-complex. The clinical phenotypic spectrum of BAFopathies is wide and involves various human neurodevelopmental disorders including syndromic and nonsyndromic intellectual disability (ID), growth retardation ^{3, 4, 5, 6, 7, 8, 9, 10}, sporadic autism ¹¹, schizophrenia ¹², and amyotrophic lateral sclerosis ¹³.

The most recognizable syndrome associated with BAF abnormalities is Coffin-Siris syndrome (CSS). This is a genetically heterogeneous ID syndrome characterized by developmental delay (DD), speech delay, coarse facial appearance, feeding

difficulties, and hypoplastic-to-absent fifth finger nails and fifth distal phalanges ¹⁴. This syndrome is associated with abnormalities in multiple subunits of the BAF-complex including the ATPase subunit SMARCA4 (MIM: 603254), the common core subunit SMARCB1 (MIM: 601607), and BAF accessory subunits such as SMARCE1/BAF57 (MIM: 603111), ARID1A (MIM: 603024), ARID1B (MIM: 614556) ¹⁵, ARID2 (MIM: 609539) ¹⁶, and DPF2 (MIM:601671) ¹⁷. CSS can result of pathogenic changes in other chromatin remodeling proteins with no direct interaction with BAF-complex, including SOX11 (MIM: 600898) ¹⁸ and PHF6 (MIM: 300414) ¹⁹.

Other BAFopathies include Nicolaides–Baraitser syndrome caused by pathogenic variants in *SMARCA2* (MIM: 600014) that has a significant phenotypic overlap with CSS and characterized by ID, sparse hair, short stature, microcephaly, brachydactyly, interphalangeal joint swellings, and epilepsy ^{20, 21}. Some individuals with clinical diagnosis of DOORS syndrome, characterized by deafness, onychodystrophy, osteodystrophy, ID and seizures, were found to carry pathogenic variants in *SMARCB1* (MIM: 601607), highlighting the clinical overlap between CSS and DOORS syndrome ²². We have previously reported mutations in another BAF subunit, *ACTL6A*, to be associated with ID ²³. Pathogenic variants in *ADNP* (MIM: 611386), encoding a transcription factor that interacts with the BAF-complex, have been identified in individuals presenting dysmorphic facial features, autism spectrum disorder (ASD), ID, hypotonia, and congenital heart defects ²⁴. Detailed phenotypic and genetic comparison between the different BAF related syndromes has been discussed elsewhere ^{4,25}.

SMARCC2 (MIM: 601734) encodes BAF170, a common core subunit of the BAF-complexes with high homology to *SMARCC1* (BAF155)²⁶. It is an intrinsic factor of glial radial cells and plays a crucial role in embryogenesis and corticogenesis, determining the mammalian body and cortical size²⁷. *Smarcc2*; *Smarcc1* double knockout mice demonstrated proteasome-mediated degradation of the entire BAF-complexes resulting in impairment of the global epigenetic and gene expression program of forebrain development²⁸. Recently, deletion of *Smarcc2* in mice revealed its role in learning and behavioral adaptation²⁹. *SMARCC2* was also reported as one of the chromatin-remodeling genes involved in ASD³⁰. Despite its significant biological role, variants in *SMARCC2* have not been directly associated with a syndrome in humans previously.



Figure 12: Pictures of ten of the individuals with *SMARCC2* variants.

Results:

We report 15 unrelated individuals (table 1 and table S1) with variants in *SMARCC2*, detected by whole exome sequencing (WES), and with clinical presentation

that includes mild to severe ID, (HP:0012736), DD with prominent speech delay (HP:0000750), behavioral abnormalities (HP:0000708), growth retardation (HP:0008897), feeding difficulties at the neonatal period (HP:0008872), hypotonia (HP:0011398), and dysmorphic features including hypertrichosis (HP:0000998), thick eyebrows (HP:0000574) / prominent supra-orbital ridges (HP:0000336), and thin upper and lower vermillion (HP:0000233) suggesting overlap with Coffin Siris and Nicolaides–Baraitser syndromes. Subjects in this cohort were gathered using GeneMatcher ³¹. All individuals' families from the different institutions agreed to participate in this study and signed appropriate consent forms. Permission for clinical photographs was given separately. Individual 4 has been reported before and was identified in a gene panel screening (1,256 genes) of 96 individuals with ID ³². The variant c.1833+2T>C in individual 10 was reported before as part of a work to identify new gene-disease associations in trio WES from 119 undiagnosed patients ³³. Individuals 1, 4, 6, 7, 9-13 and 15 had trio WES. Individuals 2, 3, 5 and 14 had a proband WES followed by Sanger confirmation for patient and parents. Biological parents of individuals 3 and father of individual 8 are not available for testing. Individual 8 had duo WES with his mother and results were re-analyzed as part of the Undiagnosed Diseases Network (UDN). Biological parents of individuals 3 and 8 are not available for testing. This individual was found to harbor an intronic splice site variant c.1833+1G>T in *SMARCC2* that was not found in his mother. Twelve of the 15 individuals have proven *de novo* variants in *SMARCC2*. Individual 2 inherited the variant from his affected father. Paternal grandparents were negative for the variant indicating a *de novo* variant in the father.

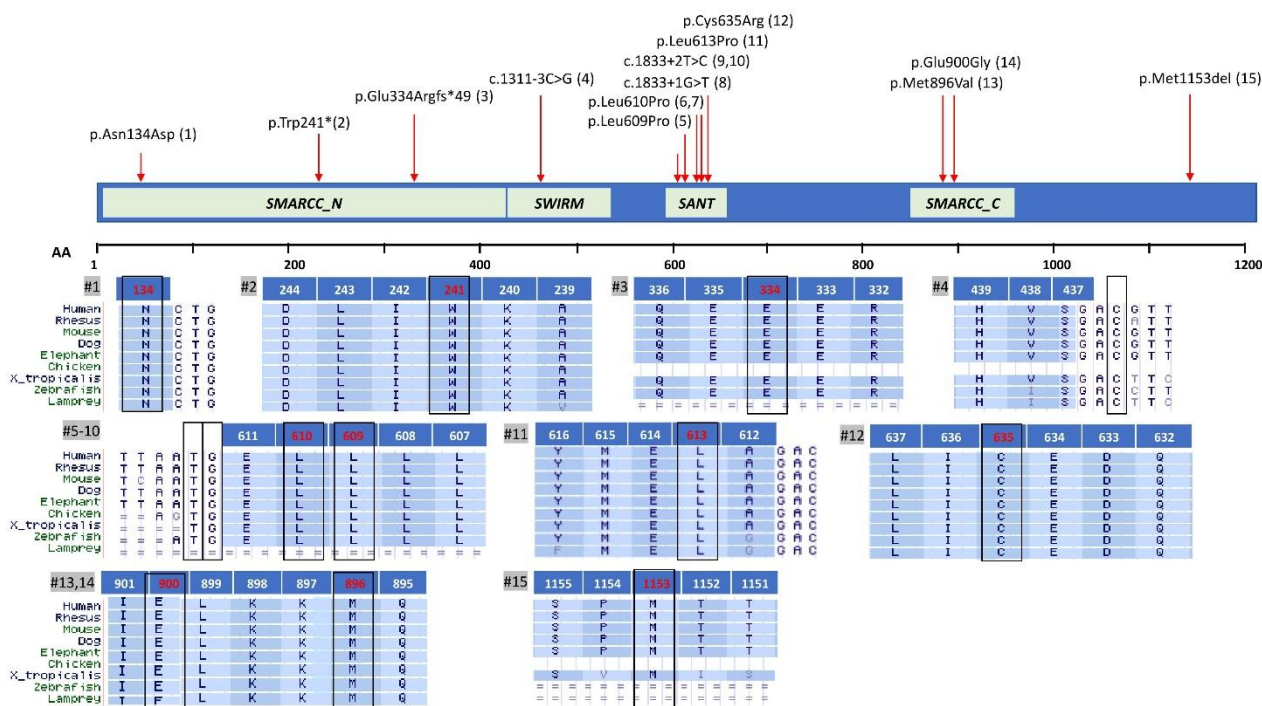


Figure 13: SMARCC2 (BAF170) main domains and variants.

Amino acid alignments demonstrate that the missense variants affect highly conserved amino acids and the variants causing splicing abnormalities affect highly conserved nucleotides (NP_003066, transcript ENST00000267064.8). The involved amino acids / nucleotides are framed in black. Number of relevant individual (#) is on gray background. For intron/exon location of the different variants see table S1.

All of the individuals presented here have some degree of ID and/or DD (table 1 and table S1). Ten (10/15, 65%) have moderate to profound DD and ID while the other five individuals have only mild ID or mild DD. 13/15 (86%) individuals have speech impairment with seven of them completely lacking language. Most individuals present muscle tone abnormalities. 13/15 (85%) individuals present significant hypotonia,

while two of the individuals present high tone or spasticity. Ten of the individuals (67%) present behavioral problems including aggression and self-injurious behavior as well as hyperactivity, hypersensitivity to touch, sleep disturbances, and obsessive and rigid behavior. Two were noted to have difficulties in social interactions, yet not qualified for formal diagnosis of autism. Eight individuals present feeding difficulties and six of them have mostly postnatal growth retardation. Individual 9 has continuous feeding difficulties with laryngomalacia and nasal feeding tube since age 14 months. 11 out of the 15 individuals are reported to have dysmorphic craniofacial features (figure 13). The most pronounced dysmorphic features are hypertrichosis (6/15), thick eyebrows / prominent supra-orbital ridges (6/15), thin upper or thick lower lip-vermilion (6/15 and 5/15), and upturned nose (6/15). Most of the individuals have normal fifth finger/toe or finger/toenail.

Eight of the 15 individuals presented here (1, 5-7, 11-14) have one of seven missense variants in *SMARCC2* (GeneBank: NM_003075.3), all are predicted to be deleterious / probably damaging according to PROVEAN and SIFT, in-silico tools to predict the functional effect of an amino acid substitution. Only individuals 2 and 3 carry truncating mutations. Individual 2 carries a heterozygous nonsense variant at position 241 (p.Trp241*), and individual 3 carries a frameshift variant (p.Glu334Argfs*49). Interestingly, both present with a mild phenotype. Individual 2 inherited the change from his affected father who presents with borderline intelligence (IQ of 72) and behavior problems. This might suggest a milder phenotype in cases with truncating mutations. All missense variants are in well-conserved amino acids in *SMARCC2* (figure 14). In vitro missense tolerance ratio (MTR) tool shows that six of the seven missense variants presented here preferentially affect one of the 25% most

intolerant residues of *SMARCC2* ($p=0.0004$) with p.Asn134Asp being the only missense variant affecting a tolerant region of this gene (figure S1). This might be consistent with the milder phenotype presented in this individual. *SMARCC2* ExAc z-score for intolerance for missense variations is significantly high (4.26) and the gene is predicted to be potentially associated with dominant conditions according to an LDA score of 2.435 by the DOMINO algorithm ³⁴. Individual 15 harbors a *de novo* in-frame deletion of methionine 1153 (c.3456_3458delCAT). Methionine 1153 deletion and its missense substitutions to various amino acids (isoleucine, threonine, valine) appear once (allele frequency 3.231×10^{-5}) and 4 times respectively in gnomAD ³⁵. This might indicate that this position is relatively tolerable for changes and explain the mild clinical presentation of individual 15. It is important to note that detailed phenotype such as IQ scores are not available for individuals included in the ExAC database.

Individuals 2 and 3 carry truncating mutations located at the N-terminus of *SMARCC2* (exons 9 and 11). Individual 2 carries a heterozygous nonsense variant (p.Trp241*), and individual 3 carries a frameshift variant (p.Glu334Argfs*49). Interestingly, both present mild phenotype. Individual 2, who presents mainly behavioral abnormalities and only mild DD, inherited the change from his affected father who presents with borderline intelligence (IQ of 72) and behavior problems. ExAC database identified only a single loss-of-function allele in *SMARCC2* (p.Glu389Aspfs*5); however, with only 11 (22%) of 49 reads at this frameshift variant site, it is unlikely to be a germline variant (Binomial exact test for 50% germline heterozygous expectation; $p = 0.0001$). The depletion of protein-truncating variants in *SMARCC2* is evident in large human population reference cohorts. Based on the ExAC cohort, *SMARCC2* achieves a LoF depletion FDR adjusted $p = 5.5 \times 10^{-11}$, ranking it

among the 1.6% most significantly LoF-depleted genes in the human exome ³⁶. It also has a probability of loss-of-function (LoF) intolerance (pLI) score of 1, supporting deleterious effect for predicted LoF pathogenic variants ³⁵. *SMARCC2* %HI score of 20.29 also predicts this gene is less likely to tolerate a loss-of-function variant or deletion ³⁷.

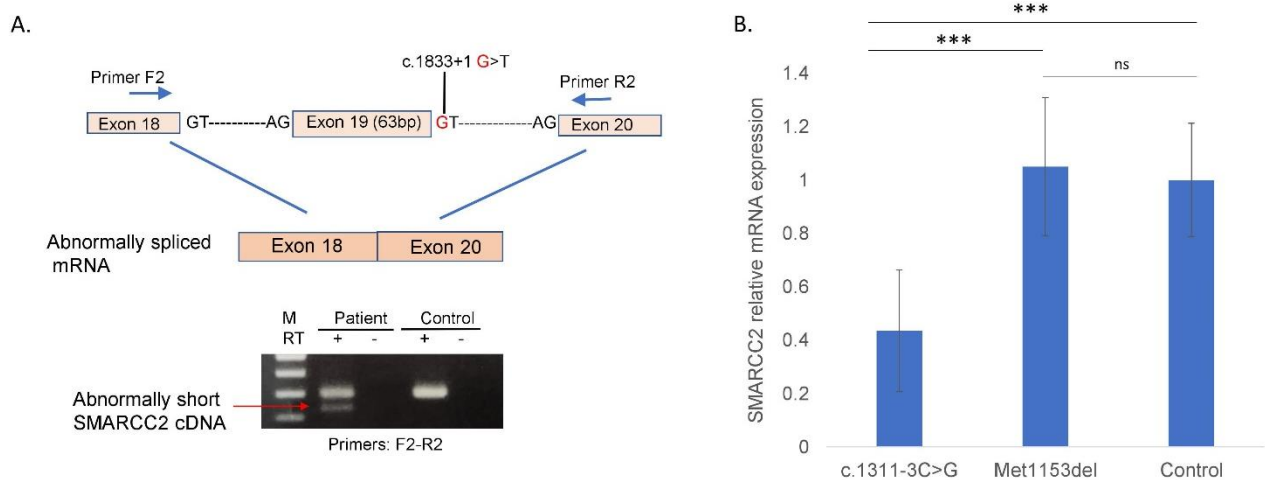


Figure 14: Splicing variant analyses in Individual 8 and 4.

A. PCR amplification of the individual's cDNA at exons 18 to 20 reveals deletion of exon 19 (amino acids 590-611) due to splicing variant c.1833+1G>T in *SMARCC2* (NM_003075). **B.** *SMARCC2* gene expression was quantified by qRT-PCR from LCLs for individuals 4 and 15 and an unrelated control, in triplicates. *SMARCC2* expression was normalized to GAPDH, and results are represented as relative expression normalized to the Control. Error bars represent SEM and significance was assessed with bilateral unpaired Student's T tests.

Four individuals harbor splicing variants. Three individuals (8-10) have one of two variants in intron 19 (c.1833+2T>C and c.1833+1G>T) that are predicted in silico to affect splicing donor site according to Human Splicing Finder (HSF). One individual

(4) presents a splicing variant in intron 14 (c.1311-3C>G) that is predicted to affect exon acceptor site. *In vitro* assay in lymphocytes from individual 8 (c.1833+1G>T) indicates this variant leads to abnormal splicing with deletion of exon 19 (amino acids 590-611) that compose the SANT domain in SMARCC2 (figure **15A**). The capture of a shorter cDNA product using RT-PCR might indicate the presence of an mRNA escaping nonsense-mediated decay (NMD). RT-qPCR quantification of *SMARCC2* in lymphocytes from individual 4 (c.1311-3C>G) demonstrate a ~50% decrease in mRNA compared to a control and a coding in-frame deletion (Individual 15, Met1153del), suggesting haploinsufficiency through NMD (figure **15B**).

Nine individuals carry missense or splicing variants in the highly conserved SWIRM (Swi3, Rsc8, and Moira) and SANT (Swi3, Ada2, NCoR and TFIIB) domains of SMARCC2 (figure **14**). The SWIRM domain interacts with DNA, binds dinucleosome structures, and mediates specific protein-protein interactions³⁸. The SANT domain has DNA-binding activity³⁹ and is believed to function as a histone tail binding module⁴⁰. Eight of the individuals reported here have a heterozygous novel variant in the SANT domain of SMARCC2. One individual has a missense variant changing leucine for proline in position 609 (p.Leu609Pro), two have a missense variant changing leucine for proline in position 610 (p.Leu610Pro), one individual has a missense variant changing leucine for proline in position 613 (p.Leu613Pro), and one has a missense variant changing cysteine to arginine in position 635 (p.Cys635Arg). Three individuals have one of two intronic variants that are predicted *in silico* to affect the same splicing donor site in SANT and one individual presents a splicing variant at the SWIRM domain. This group of nine individuals present with moderate to severe DD and ID and with severe speech impairment, with six of them having no language at

all. Five out of these nine individuals (55%) have abnormal brain MRI findings including small corpus callosum and generalized cerebral atrophy. Five out of the nine present significant growth retardation and four of them have seizure disorder. It seems that individuals carrying a pathogenic variant in SWIRM or SANT domains have a more severe presentation. Interestingly, only the two individuals harboring the p.Leu610Pro variant have cardiovascular abnormalities. An echocardiogram for individual 6 revealed distended left coronary artery and for individual 7 mild non-progressive dilatation of the ascending aorta (Z score of 2.54).

To further investigate the pathogenicity of the variants we calculated the geometric mean distance between the observed mutations to assay the clustering of the mutations ⁴¹. As this calculation only considers the length of the spliced transcript, intronic splicing mutations were annotated according to the closest coding nucleotide. Compared to ten million random permutations of mutations along the transcript, the observed mutations are significantly clustered (p-value 10^{-7}), suggesting a pathomechanism which does not operate through haploinsufficiency alone. Interestingly, removing the frameshift and nonsense mutations from the analysis does not affect the significance, but taking out the splicing mutations does (p-value 0.12). This result points towards a more important role in the pathogenicity for the SWIRM and SANT domains than for the SMARCC_N region, and a corollary of this interpretation is the hypothesis that expressed but incorrectly-spliced transcripts could function through a dominant negative mechanism.

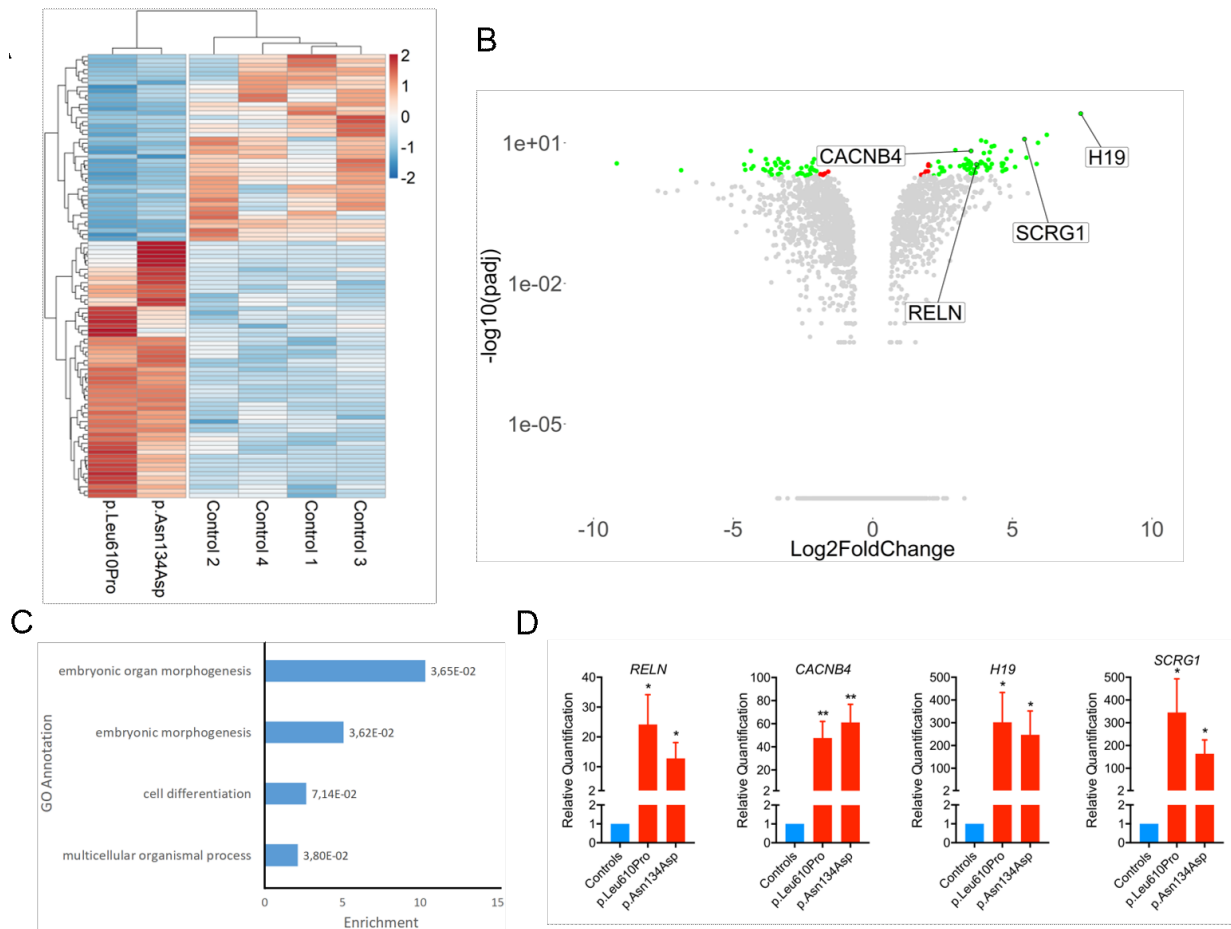


Figure 15: Fibroblasts harboring p.Leu610Pro or p.Asn134Asp SMARCC2 mutation have differential gene expression patterns when compared to fibroblasts from healthy controls.

RNAseq was performed on four healthy controls and two patient fibroblasts to assay gene expression. (A) Heatmap of Differentially Expressed Genes (DEGs). Log₂ Fold Change (Log₂FC) and p-values adjusted for 10% False Discovery Rate (padj) were calculated for the two individuals and four controls respectively pooled together using the “DESeq2” R package. Normalized count values from “DESeq2” were plotted and scaled row-wise using the “pheatmap” R package for genes with a padj lower than 0.01 and an absolute Log₂FC value higher than 2. DEGs numbers are shown as Venn diagrams with each individual compared separately to controls using the “DESeq2” R package, for up-regulated genes (B) and down-regulated genes (C).

(D) Volcano plot showing differentially regulated genes as fold-change versus adjusted p-values. Log_2FC and $-\log_{10}(\text{padj})$ values were plotted using the “ggplot2” R package as a volcano plot. Red dots represent genes with a padj lower than 0.01 and a $\text{abs}(\text{Log}_2\text{FC})$ lower than 2. Green dots represent genes with a padj lower than 0.01 and a $\text{abs}(\text{Log}_2\text{FC})$ higher than 2 (E) Histogram of the enriched GO annotations with FDR adjusted p-values lower than 0.01 calculated using the GOrilla web application using DEGs for the two individuals pooled together. (F) RT-qPCR analysis of

To better understand the impact of *SMARCC2* mutations on gene-expression, we performed an RNAseq analysis on fibroblasts from individual 1 with p.Asn134Asp mutation and individual 7 with p.Leu610Pro mutation. We analyzed both samples together to assess shared differential gene expression patterns. Analysis of Differentially Expressed Genes (DEGs) shows an almost equal amount of DEGs upregulated and downregulated in these two individuals (figure 16A). Gene Ontology (GO) analysis of these DEGs using the GOrilla web application demonstrates a significant enrichment in genes related to embryonic morphogenesis, multicellular organismal process, and developmental process (figure 16E). We also looked at DEGs specific to each individual. Consistent with the milder phenotype of individual 1, the p.Asn134Asp mutation seems to affect the expression of fewer genes, as this individual only had 25 DEGs (table S3) compared to the 59 of individual 7 (table S4 and figure 4B and 4C). Of those, they shared 12 DEGs. Moreover, many genes that were not significantly differentially expressed in individual 1, but were in individual 7, did show a tendency to have relative changes in expression in the same direction. Consistent with *in silico* predictions, this suggests that the p.Asn134Asp mutation, located at the N-

terminus of the protein has a milder effect than the p.Leu610Pro mutation that is located in the SANT domain. Among the shared DEGs we identified genes with possible role in regulation of neuronal development and function, namely *H19*, *SCRG1*, *RELN* and *CACNB4*.

H19 codes for a long noncoding RNA and is among the most upregulated DEGs ($\log_2FC=7.4$, $padj=7.18E-43$) in both individuals compared to controls (table S2). *H19* is an imprinted gene and has been implicated in the regulation of growth and embryonic development, tumorigenesis and epilepsy-induced astrocyte and microglia activation ^{42, 43, 44}. Interestingly, knocking down *SMARCA4* in MCF10A epithelial cell line also results in significant *H19* upregulation ¹ suggesting that *H19* expression is affected by BAF complex modification and that *SMARCC2* p.Asn134Asp and p.Leu610Pro mutations could indeed affect the complex integrity and/or function.

SCRG1 is another significant DEG ($\log_2FC=5.43$, $padj=9.30E-13$) mainly expressed in the human brain, and its level is highly regulated during postnatal development, being absent in the fetal brain ⁴⁵. Duplication of the gene has been suggested to contribute to ID ⁴⁶. In human mesenchymal stem cells, *SCRG1* expression is important for the maintenance of self-renewal and preventing osteoblastic differentiation⁴⁷. Interestingly, impaired postnatal radial glial cells self-renewal was observed in conditional *Smarcc2*-deficient animals ²⁹, and abolishing the expression of *Smarca4* or *Smarcc2* in postnatal neuronal stem cells results in preferential onset of gliogenesis instead of neurogenesis ^{48,29}.

Another overexpressed gene is *RELN* ($\log_2FC=3.72$, $padj=3.39E-04$), which encodes Reelin, a large secreted matrix serine protease that is involved in layering of neurons in the cortex and cerebellum. It also regulates microtubule function in neurons and neuronal migration, and its enzymatic activity is important for the modulation of cell adhesion. Recessive *RELN* mutations cause lissencephaly (MIM: 257320), while dominant mutations cause epilepsy (MIM: 616436) ⁴⁹.

Finally, *CACNB4*, which also has significantly increased expression ($\log_2FC=3.52$, $padj=1.88E-07$), encodes a beta auxiliary subunit of voltage-gated calcium ions channels. Calcium channels play an essential role in the nervous system for neurotransmitter release as well as the regulation of gene expression ^{50, 51}. Interestingly, *CACNB4* was shown to translocate to the nucleus to regulate gene expression ^{52, 53}. Different isoforms of the β -subunit of voltage-gated calcium ions channels have been shown to control the transcription of genes by recruiting proteins involved in DNA remodeling such as the heterochromatin protein 1 ⁵⁴. Also, the β -subunit was reported to downregulate Wnt signaling, a crucial player in neural development ⁵⁵, through interaction with the transcription factor TCF4 which is associated with Pitt-Hopkins syndrome (MIM: 610954) ⁵³. Mutations in other subunits of voltage-gated channels have been associated with mental disorders, autism spectrum disorders (ASD) and cardiovascular problems ^{56, 57, 58, 59}. In humans, *CACNB4* mutations have been associated with episodic ataxia (MIM: 613855), juvenile myoclonic and idiopathic generalized epilepsy (MIM: 607682).

To validate the relevance of our results and to understand the low overlap between DEGs in the two assessed individuals, we compared our results to other public

datasets involving BAF-complex components. We looked at datasets in both human and mouse cell lines and primary cells in culture (figure S2) ^{60, 1, 61, 62, 63}. The number of DEGs varies significantly between assays from just 18 to 3038. We calculated and plotted the overlapping genes for each dataset, showing varying degrees of overlap (0% to 83%) suggesting a great heterogeneity in genes affected by BAF-complex defects, in both a subunit and a cell type specific manner. Thus, we hypothesize that pathogenic mutations in different domains of the same subunit, as seen in the two individuals assessed in our study, could lead to different gene expression patterns as they could affect the interaction with unique subunits or transcription factors and could explain the low overlap we obtained. Still, 48% of the DEGs in *SMARCC2* individuals' fibroblasts are found in at least one other dataset, suggesting some commonalities, and confirming a general BAF-complex-linked transcriptomic profile.

Discussion:

Spatiotemporal regulation of the different subunits assembly and the activity of the BAF-complex are well orchestrated and essential to enable normal development and proper functioning of the nervous system ^{2, 64, 65}. With the exception of impaired cognitive function observed in *Smarcc2*-deficient mice, no mutations in *SMARCC2* have been associated with a particular phenotype/disease. Here we show that mutations in *SMARCC2* affect gene expression, and in particular a subset of putative genes potentially implicated in neurogenesis and proper functioning of the nervous system, namely *H19*, *SCRG1*, *RELN* and *CACNB4*. Our results further suggest that mutations of *SMARCC2* may regulate pathways important for postnatal gliogenesis, differentiation and function of astrocytes and oligodendrocytes. Further studies need to

be undertaken in other models to determine the role of these genes in individuals affected by *SMARCC2* mutations.

Neurodevelopmental disorders associated with pathogenic variants in BAF chromatin remodeling complex subunits present overlapping clinical phenotype. Comparing the clinical presentation of the different associated conditions (table 2 ^{66, 4, 3, 67, 68, 69, 23}), emphasizes the wide clinical spectrum of the disorders and the importance of trio WES in the diagnostic process. Other than our observation that missense and splicing variants in the highly conserved SANT and SWIRM domains uniformly lead to a severe phenotype, we could not point out a clear genotype-phenotype correlation. Missense variants outside of these two domains were found mostly but not exclusively in patients with mild presentation suggesting variable degree of effect on BAF-complex function for missense variants. Pathogenic variants reported in SMARC subunits (*SMARCA2*, *SMARCA4*, *SMARCB1* and *SMARCE1*) are mostly missense and in-frame deletions (table 2) with dominant negative effect mechanism causing Coffin-Siris and CS-like syndromes ^{8, 25}. Recently, mutations identified in BAF-subunit DPF2 were suggested to cause CSS-like phenotype in a dominant negative mechanism related both to missense and splicing / truncating variants through NMD escape ¹⁷.

Given that *SMARCC1* and *SMARCC2* are paralogous genes, that they can form heterodimers or homodimers ⁷⁰, and that both share functional scaffolding properties ²⁸, it is conceivable that *SMARCC1* could at least partially compensate for the loss of *SMARCC2* leading to a milder phenotype in clinically milder cases with truncating mutations in the SMARCC N-terminal region. Nevertheless, our results also suggest that the mutant mRNA of individual 4 with a severe phenotype (c.1311-3C>G)

undergoes NMD as one of the consequences of the splice site mutation, therefore haploinsufficiency can also be seen in individuals with a severe phenotype. As neither RNA nor cells were available for these individuals with N-terminal truncating mutation, we cannot exclude the possibility that the mRNA could escape NMD and the truncated protein remained functional to perform essential function or to allow for a compensatory mechanism to take place. Finally, different SMARCC2 isoforms could be expressed in different tissues or at different times during development, thus the location of truncating mutations could affect the phenotype since they may affect only certain isoforms. Indeed, if there are N-terminal truncated isoforms of SMARCC2 expressed specifically in the nervous system, the N-terminal truncation mutations causing NMD would have no impact on that isoform while a mutation in the SWIRM domain would. Such specific alternative splicing could explain the discrepancy of mild and severe phenotype being both associated with haploinsufficiency. Interestingly, Kazantseva and colleagues have shown that N-terminally truncated isoforms of SMARCE1 (BAF57) are exclusively expressed in neuronal tissues and contribute to the diversity of the BAF complex in neurons ⁷¹. A similar mechanism could occur for SMARCC2, and further research into its neuronal expression would be needed to answer this question. The wide phenotypic variety and the lack of clear genotype-phenotype correlation of the conditions associated with BAF-complex subunits were also suggested to be attributed to a possible dosage dependent manner of the ATP-dependent chromatin remodeling machinery ²⁵. This, with the significant intolerance of the BAF-complex subunits genes to LOF variants, might support a contribution of LOF variants to the phenotype.

The *SMARCC2*-related condition presented here overlaps with other BAFopathies suggesting a CSS and Nicolaides-Baraitser-like syndrome characterized by intellectual disability, developmental delay with significant speech delay and behavioral abnormalities. We provide evidence of dysregulated expression of *H19*, *SCRGI*, *RELN* and *CACNB4*, the relevance of which will need to be explored in future model organism studies.

Consortia:

The members of the Undiagnosed Diseases Network are as follows:

David R. Adams, Mercedes E. Alejandro, Patrick Allard, Mahshid S. Azamian, Carlos A. Bacino, Ashok Balasubramanyam, Hayk Barseghyan, Gabriel F. Batzli, Alan H. Beggs, Babak Behnam, Anna Bican, David P. Bick, Camille L. Birch, Devon Bonner, Braden E. Boone, Bret L. Bostwick, Lauren C. Briere, Donna M. Brown, Matthew Brush, Elizabeth A. Burke, Lindsay C. Burrage, Shan Chen, Gary D. Clark, Terra R. Coakley, Joy D. Cogan, Cynthia M. Cooper, Heidi Cope, William J. Craigen, Precilla D'Souza, Mariska Davids, Jyoti G. Dayal, Esteban C. Dell'Angelica, Shweta U. Dhar, Ani Dillon, Katrina M. Dipple, Laurel A. Donnell-Fink, Naghmeh Dorrani, Daniel C. Dorset, Emilie D. Douine, David D. Draper, David J. Eckstein, Lisa T. Emrick, Christine M. Eng, Ascia Eskin, Cecilia Esteves, Tyra Estwick, Carlos Ferreira, Brent L. Fogel, Noah D. Friedman, William A. Gahl, Emily Glanton, Rena A. Godfrey, David B. Goldstein, Sarah E. Gould, Jean-Philippe F. Gourdine, Catherine A. Groden, Andrea L. Gropman, Melissa Haendel, Rizwan Hamid, Neil A. Hanchard, Lori H. Handley, Matthew R. Herzog, Ingrid A. Holm, Jason Hom, Ellen M. Howerton, Yong Huang, Howard J. Jacob, Mahim Jain, Yonghui Jiang, Jean M. Johnston, Angela L. Jones, Isaac S. Kohane, Donna M.

Krasnewich, Elizabeth L. Krieg, Joel B. Krier, Seema R. Lalani, C. Christopher Lau, Jozef Lazar, Brendan H. Lee, Hane Lee, Shawn E. Levy, Richard A. Lewis, Sharyn A. Lincoln, Allen Lipson, Sandra K. Loo, Joseph Loscalzo, Richard L. Maas, Ellen F. Macnamara, Calum A. MacRae, Valerie V. Maduro, Marta M. Majcherska, May Christine V. Malicdan, Laura A. Mamounas, Teri A. Manolio, Thomas C. Markello, Ronit Marom, Julian A. Martínez-Agosto, Shruti Marwaha, Thomas May, Allyn McConkie-Rosell, Colleen E. McCormack, Alexa T. McCray, Matthew Might, Paolo M. Moretti, Marie Morimoto, John J. Mulvihill, Jennifer L. Murphy, Donna M. Muzny, Michele E. Nehrebecky, Stan F. Nelson, J. Scott Newberry, John H. Newman, Sarah K. Nicholas, Donna Novacic, Jordan S. Orange, J. Carl Pallais, Christina G.S. Palmer, Jeanette C. Papp, Neil H. Parker, Loren D.M. Pena, John A. Phillips III, Jennifer E. Posey, John H. Postlethwait, Lorraine Potocki, Barbara N. Pusey, Chloe M. Reuter, Amy K. Robertson, Lance H. Rodan, Jill A. Rosenfeld, Jacinda B. Sampson, Susan L. Samson, Kelly Schoch, Molly C. Schroeder, Daryl A. Scott, Prashant Sharma, Vandana Shashi, Rebecca Signer, Edwin K. Silverman, Janet S. Sinsheimer, Kevin S. Smith, Rebecca C. Spillmann, Kimberly Splinter, Joan M. Stoler, Nicholas Stong, Jennifer A. Sullivan, David A. Sweetser, Cynthia J. Tifft, Camilo Toro, Alyssa A. Tran, Tiina K. Urv, Zaheer M. Valivullah, Eric Vilain, Tiphane P. Vogel, Colleen E. Wahl, Nicole M. Walley, Chris A. Walsh, Patricia A. Ward, Katrina M. Waters, Monte Westerfield, Anastasia L. Wise, Lynne A. Wolfe, Elizabeth A. Worthey, Shinya Yamamoto, Yaping Yang, Guoyun Yu, Diane B. Zastrow, and Allison Zheng.

Conflicts of Interest

The Department of Molecular and Human Genetics at Baylor College of Medicine receives revenue from clinical genetic testing performed at Baylor Genetics.

Megan Cho and Heather McLaughlin are employees of GeneDx, Inc., a wholly-owned subsidiary of OPKO Health, Inc. All other authors declare no competing interests.

Acknowledgments

We thank the patients and their families for their participation in this study. We thank the CIHR and FRSQ for clinician-scientist awards to PMC. Spanish patients study was supported by grant PI14/00350 (Instituto de Salud Carlos III -Acción Estratégica en Salud 2013–2016; FEDER - Fondo Europeo de Desarrollo Regional). Partial Research reported in this manuscript was supported by the NIH Common Fund, through the Office of Strategic Coordination/Office of the NIH Director under Award Number(s) U01HG007672 (Duke University) and U01HG007942 (Baylor College of Medicine- Sequencing). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health and partial funding for this study was provided by UCB Celltech.

Web Resources

ExAC Browser, <http://exac.broadinstitute.org/>

GeneBank, <https://www.ncbi.nlm.nih.gov/genbank/>

GeneMatcher, <https://www.genematcher.org/>

GnomAD Browser, <http://gnomad.broadinstitute.org/>

Genome Browser, <https://genome.ucsc.edu/>

Human Splicing Finder, <http://www.umd.be/HSF3/>

Human Phenotype Ontology (HPO), <http://human-phenotype-ontology.org/>

OMIM, <https://www.omim.org/>

Polyphen-2, <http://genetics.bwh.harvard.edu/pph2>

PROVEAN, <http://provean.jcvi.org/>

SIFT, <http://sift.jcvi.org/>

Genetic Intolerance, <http://genic-intolerance.org/>

Missense Tolerance Ratio (MTR) tool, <http://mtr-viewer.mdhs.unimelb.edu.au/mtr-viewer/>

REFERENCES

1. Barutcu, A.R., Lajoie, B.R., Fritz, A.J., McCord, R.P., Nickerson, J.A., van Wijnen, A.J., Lian, J.B., Stein, J.L., Dekker, J., Stein, G.S., et al. (2016). SMARCA4 regulates gene expression and higher-order chromatin structure in proliferating mammary epithelial cells. *Genome Res* 26, 1188-1201.
2. Ronan, J.L., Wu, W., and Crabtree, G.R. (2013). From neural development to cognition: unexpected roles for chromatin. *Nat Rev Genet* 14, 347-359.
3. Santen, G.W., Aten, E., Vulto-van Silfhout, A.T., Pottinger, C., van Bon, B.W., van Minderhout, I.J., Snowdowne, R., van der Lans, C.A., Boogaard, M., Linssen, M.M., et al. (2013). Coffin-Siris syndrome and the BAF complex: genotype-phenotype study in 63 patients. *Hum Mutat* 34, 1519-1528.
4. Koshi, T., Okamoto, N., and Coffin-Siris Syndrome International, C. (2014). Genotype-phenotype correlation of Coffin-Siris syndrome caused by mutations in SMARCB1, SMARCA4, SMARCE1, and ARID1A. *Am J Med Genet C Semin Med Genet* 166C, 262-275.
5. Sousa, S.B., Hennekam, R.C., and Nicolaides-Baraitser Syndrome International, C. (2014). Phenotype and genotype in Nicolaides-Baraitser syndrome. *Am J Med Genet C Semin Med Genet* 166C, 302-314.
6. Hoyer, J., Ekici, A.B., Ende, S., Popp, B., Zweier, C., Wiesener, A., Wohlleber, E., Dufke, A., Rossier, E., Petsch, C., et al. (2012). Haploinsufficiency of ARID1B, a member of the SWI/SNF-a chromatin-remodeling complex, is a frequent cause of intellectual disability. *Am J Hum Genet* 90, 565-572.
7. Halgren, C., Kjaergaard, S., Bak, M., Hansen, C., El-Schich, Z., Anderson, C.M., Henriksen, K.F., Hjalgrim, H., Kirchhoff, M., Bijlsma, E.K., et al. (2012). Corpus callosum abnormalities, intellectual disability, speech impairment, and autism in patients with haploinsufficiency of ARID1B. *Clin Genet* 82, 248-255.
8. Van Houdt, J.K., Nowakowska, B.A., Sousa, S.B., van Schaik, B.D., Seuntjens, E., Avonce, N., Sifrim, A., Abdul-Rahman, O.A., van den Boogaard, M.J., Bottani, A., et al. (2012). Heterozygous missense mutations in SMARCA2 cause Nicolaides-Baraitser syndrome. *Nat Genet* 44, 445-449, S441.

9. Vandeweyer, G., Helsmoortel, C., Van Dijck, A., Vulto-van Silfhout, A.T., Coe, B.P., Bernier, R., Gerdts, J., Rooms, L., van den Ende, J., Bakshi, M., et al. (2014). The transcriptional regulator ADNP links the BAF (SWI/SNF) complexes with autism. *Am J Med Genet C Semin Med Genet* 166C, 315-326.
10. Santen, G.W., Kriek, M., and van Attikum, H. (2012). SWI/SNF complex in disorder: SWItching from malignancies to intellectual disability. *Epigenetics* 7, 1219-1224.
11. Neale, B.M., Kou, Y., Liu, L., Ma'ayan, A., Samocha, K.E., Sabo, A., Lin, C.F., Stevens, C., Wang, L.S., Makarov, V., et al. (2012). Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* 485, 242-245.
12. Loe-Mie, Y., Lepagnol-Bestel, A.M., Maussion, G., Doron-Faigenboim, A., Imbeaud, S., Delacroix, H., Aggerbeck, L., Pupko, T., Gorwood, P., Simonneau, M., et al. (2010). SMARCA2 and other genome-wide supported schizophrenia-associated genes: regulation by REST/NRSF, network organization and primate-specific evolution. *Hum Mol Genet* 19, 2841-2857.
13. Chesi, A., Staahl, B.T., Jovicic, A., Couthouis, J., Fasolino, M., Raphael, A.R., Yamazaki, T., Elias, L., Polak, M., Kelly, C., et al. (2013). Exome sequencing to identify de novo mutations in sporadic ALS trios. *Nat Neurosci* 16, 851-855.
14. Schrier, S.A., Bodurtha, J.N., Burton, B., Chudley, A.E., Chiong, M.A., D'Avanzo M, G., Lynch, S.A., Musio, A., Nyazov, D.M., Sanchez-Lara, P.A., et al. (2012). The Coffin-Siris syndrome: a proposed diagnostic approach and assessment of 15 overlapping cases. *Am J Med Genet A* 158A, 1865-1876.
15. Tsurusaki, Y., Okamoto, N., Ohashi, H., Kosho, T., Imai, Y., Hibi-Ko, Y., Kaname, T., Naritomi, K., Kawame, H., Wakui, K., et al. (2012). Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nat Genet* 44, 376-378.
16. Bramswig, N.C., Caluseriu, O., Ludecke, H.J., Bolduc, F.V., Noel, N.C., Wieland, T., Surowy, H.M., Christen, H.J., Engels, H., Strom, T.M., et al. (2017). Heterozygosity for ARID2 loss-of-function mutations in individuals with a Coffin-Siris syndrome-like phenotype. *Hum Genet* 136, 297-305.
17. Vasileiou, G., Vergarajauregui, S., Endeley, S., Popp, B., Buttner, C., Ekici, A.B., Gerard, M., Bramswig, N.C., Albrecht, B., Clayton-Smith, J., et al. (2018). Mutations in the BAF-Complex Subunit DPF2 Are Associated with Coffin-Siris Syndrome. *Am J Hum Genet* 102, 468-479.
18. Hempel, A., Pagnamenta, A.T., Blyth, M., Mansour, S., McConnell, V., Kou, I., Ikegawa, S., Tsurusaki, Y., Matsumoto, N., Lo-Castro, A., et al. (2016). Deletions and de novo mutations of SOX11 are associated with a neurodevelopmental disorder with features of Coffin-Siris syndrome. *J Med Genet* 53, 152-162.
19. Zweier, C., Rittinger, O., Bader, I., Berland, S., Cole, T., Degenhardt, F., Di Donato, N., Graul-Neumann, L., Hoyer, J., Lynch, S.A., et al. (2014). Females with de novo aberrations in

- PHF6: clinical overlap of Borjeson-Forssman-Lehmann with Coffin-Siris syndrome. *Am J Med Genet C Semin Med Genet* 166C, 290-301.
20. Nicolaides, P., and Baraitser, M. (1993). An unusual syndrome with mental retardation and sparse hair. *Clin Dysmorphol* 2, 232-236.
 21. Preteggiani, E., Mari, F., Renieri, A., Penco, S., and Dotti, M.T. (2016). Nicolaides-Baraitser syndrome: defining a phenotype. *J Neurol* 263, 1659-1660.
 22. Campeau, P.M., Hennekam, R.C., and group, D.s.c. (2014). DOORS syndrome: phenotype, genotype and comparison with Coffin-Siris syndrome. *Am J Med Genet C Semin Med Genet* 166C, 327-332.
 23. Marom, R., Jain, M., Burrage, L.C., Song, I.W., Graham, B.H., Brown, C.W., Stevens, S.J.C., Stegmann, A.P.A., Gunter, A.T., Kaplan, J.D., et al. (2017). Heterozygous variants in ACTL6A, encoding a component of the BAF complex, are associated with intellectual disability. *Hum Mutat* 38, 1365-1371.
 24. Helsmoortel, C., Vulto-van Silfhout, A.T., Coe, B.P., Vandeweyer, G., Rooms, L., van den Ende, J., Schuurs-Hoeijmakers, J.H., Marcelis, C.L., Willemsen, M.H., Vissers, L.E., et al. (2014). A SWI/SNF-related autism syndrome caused by de novo mutations in ADNP. *Nat Genet* 46, 380-384.
 25. Bogershausen, N., and Wollnik, B. (2018). Mutational Landscapes and Phenotypic Spectrum of SWI/SNF-Related Intellectual Disability Disorders. *Front Mol Neurosci* 11, 252.
 26. Ring, H.Z., Vameghi-Meyers, V., Wang, W., Crabtree, G.R., and Francke, U. (1998). Five SWI/SNF-related, matrix-associated, actin-dependent regulator of chromatin (SMARC) genes are dispersed in the human genome. *Genomics* 51, 140-143.
 27. Tuoc, T.C., Boretius, S., Sansom, S.N., Pitulescu, M.E., Frahm, J., Livesey, F.J., and Stoykova, A. (2013). Chromatin regulation by BAF170 controls cerebral cortical size and thickness. *Dev Cell* 25, 256-269.
 28. Narayanan, R., Pirouz, M., Kerimoglu, C., Pham, L., Wagener, R.J., Kiszka, K.A., Rosenbusch, J., Seong, R.H., Kessel, M., Fischer, A., et al. (2015). Loss of BAF (mSWI/SNF) Complexes Causes Global Transcriptional and Chromatin State Changes in Forebrain Development. *Cell Rep* 13, 1842-1854.
 29. Tuoc, T., Dere, E., Radyushkin, K., Pham, L., Nguyen, H., Tonchev, A.B., Sun, G., Ronnenberg, A., Shi, Y., Staiger, J.F., et al. (2017). Ablation of BAF170 in Developing and Postnatal Dentate Gyrus Affects Neural Stem Cell Proliferation, Differentiation, and Learning. *Mol Neurobiol* 54, 4618-4635.
 30. Ben-David, E., and Shifman, S. (2013). Combined analysis of exome sequencing points toward a major role for transcription regulation during brain development in autism. *Mol Psychiatry* 18, 1054-1056.

31. Sobreira, N., Schiettecatte, F., Valle, D., and Hamosh, A. (2015). GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. *Hum Mutat* 36, 928-930.
32. Martinez, F., Caro-Llopis, A., Rosello, M., Oltra, S., Mayo, S., Monfort, S., and Orellana, C. (2017). High diagnostic yield of syndromic intellectual disability by targeted next-generation sequencing. *J Med Genet* 54, 87-92.
33. Zhu, X., Petrovski, S., Xie, P., Ruzzo, E.K., Lu, Y.F., McSweeney, K.M., Ben-Zeev, B., Nissenkorn, A., Anikster, Y., Oz-Levi, D., et al. (2015). Whole-exome sequencing in undiagnosed genetic diseases: interpreting 119 trios. *Genet Med* 17, 774-781.
34. Quinodoz, M., Royer-Bertrand, B., Cisarova, K., Di Gioia, S.A., Superti-Furga, A., and Rivolta, C. (2017). DOMINO: Using Machine Learning to Predict Genes Associated with Dominant Disorders. *Am J Hum Genet* 101, 623-629.
35. Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536, 285-291.
36. Petrovski, S., Gussow, A.B., Wang, Q., Halvorsen, M., Han, Y., Weir, W.H., Allen, A.S., and Goldstein, D.B. (2015). The Intolerance of Regulatory Sequence to Genetic Variation Predicts Gene Dosage Sensitivity. *PLoS Genet* 11, e1005492.
37. Huang, N., Lee, I., Marcotte, E.M., and Hurles, M.E. (2010). Characterising and predicting haploinsufficiency in the human genome. *PLoS Genet* 6, e1001154.
38. Aravind, L., and Iyer, L.M. (2002). The SWIRM domain: a conserved module found in chromosomal proteins points to novel chromatin-modifying activities. *Genome Biol* 3, RESEARCH0039.
39. Yoneyama, M., Tochio, N., Umehara, T., Koshiba, S., Inoue, M., Yabuki, T., Aoki, M., Seki, E., Matsuda, T., Watanabe, S., et al. (2007). Structural and functional differences of SWIRM domain subtypes. *J Mol Biol* 369, 222-238.
40. Boyer, L.A., Latek, R.R., and Peterson, C.L. (2004). The SANT domain: a unique histone-tail-binding module? *Nat Rev Mol Cell Biol* 5, 158-163.
41. Lelieveld, S.H., Wiel, L., Venselaar, H., Pfundt, R., Vriend, G., Veltman, J.A., Brunner, H.G., Vissers, L., and Gilissen, C. (2017). Spatial Clustering of de Novo Missense Mutations Identifies Candidate Neurodevelopmental Disorder-Associated Genes. *Am J Hum Genet* 101, 478-484.
42. Gabory, A., Jammes, H., and Dandolo, L. (2010). The H19 locus: role of an imprinted non-coding RNA in growth and development. *Bioessays* 32, 473-480.
43. Raveh, E., Matouk, I.J., Gilon, M., and Hochberg, A. (2015). The H19 Long non-coding RNA in cancer initiation, progression and metastasis - a proposed unifying theory. *Mol Cancer* 14, 184.

44. Han, C.L., Ge, M., Liu, Y.P., Zhao, X.M., Wang, K.L., Chen, N., Meng, W.J., Hu, W., Zhang, J.G., Li, L., et al. (2018). LncRNA H19 contributes to hippocampal glial cell activation via JAK/STAT signaling in a rat model of temporal lobe epilepsy. *J Neuroinflammation* 15, 103.
45. Dron, M., Dandoy-Dron, F., Guillo, F., Benboudjema, L., Hauw, J.J., Lebon, P., Dormont, D., and Tovey, M.G. (1998). Characterization of the human analogue of a Scrapie-responsive gene. *J Biol Chem* 273, 18015-18018.
46. Kashevarova, A.A., Nazarenko, L.P., Skryabin, N.A., Salyukova, O.A., Chechetkina, N.N., Tolmacheva, E.N., Sazhenova, E.A., Magini, P., Graziano, C., Romeo, G., et al. (2014). Array CGH analysis of a cohort of Russian patients with intellectual disability. *Gene* 536, 145-150.
47. Aomatsu, E., Takahashi, N., Sawada, S., Okubo, N., Hasegawa, T., Taira, M., Miura, H., Ishisaki, A., and Chosa, N. (2014). Novel SCRG1/BST1 axis regulates self-renewal, migration, and osteogenic differentiation potential in mesenchymal stem cells. *Sci Rep* 4, 3652.
48. Ninkovic, J., Steiner-Mezzadri, A., Jawerka, M., Akinci, U., Masserdotti, G., Petricca, S., Fischer, J., von Holst, A., Beckers, J., Lie, C.D., et al. (2013). The BAF complex interacts with Pax6 in adult neural progenitors to establish a neurogenic cross-regulatory transcriptional network. *Cell Stem Cell* 13, 403-418.
49. Wasser, C.R., and Herz, J. (2017). Reelin: Neurodevelopmental Architect and Homeostatic Regulator of Excitatory Synapses. *J Biol Chem* 292, 1330-1338.
50. Mochida, S. (2018). Presynaptic calcium channels. *Neurosci Res* 127, 33-44.
51. Greer, P.L., and Greenberg, M.E. (2008). From synapse to nucleus: calcium-dependent gene transcription in the control of synapse development and function. *Neuron* 59, 846-860.
52. Ronjat, M., Kiyonaka, S., Barbado, M., De Waard, M., and Mori, Y. (2013). Nuclear life of the voltage-gated Cacnb4 subunit and its role in gene transcription regulation. *Channels (Austin)* 7, 119-125.
53. Rima, M., Daghsni, M., Lopez, A., Fajloun, Z., Lefrancois, L., Dunach, M., Mori, Y., Merle, P., Bruses, J.L., De Waard, M., et al. (2017). Down-regulation of the Wnt/beta-catenine signaling pathway by Cacnb4. *Mol Biol Cell* 28, 3699-3708.
54. Tadmouri, A., Kiyonaka, S., Barbado, M., Rousset, M., Fablet, K., Sawamura, S., Bahembera, E., Pernet-Gallay, K., Arnoult, C., Miki, T., et al. (2012). Cacnb4 directly couples electrical activity to gene expression, a process defective in juvenile epilepsy. *EMBO J* 31, 3730-3744.
55. Mulligan, K.A., and Cheyette, B.N. (2012). Wnt signaling in vertebrate neural development and function. *J Neuroimmune Pharmacol* 7, 774-787.

56. Pinggera, A., and Striessnig, J. (2016). Cav 1.3 (CACNA1D) L-type Ca(2+) channel dysfunction in CNS disorders. *J Physiol* 594, 5839-5849.
57. Zamponi, G.W. (2016). Targeting voltage-gated calcium channels in neurological and psychiatric diseases. *Nat Rev Drug Discov* 15, 19-34.
58. Breitenkamp, A.F., Matthes, J., and Herzig, S. (2015). Voltage-gated Calcium Channels and Autism Spectrum Disorders. *Curr Mol Pharmacol* 8, 123-132.
59. Soldatov, N.M. (2015). CACNB2: An Emerging Pharmacological Target for Hypertension, Heart Failure, Arrhythmia and Mental Disorders. *Curr Mol Pharmacol* 8, 32-42.
60. Wang, X., Lee, R.S., Alver, B.H., Haswell, J.R., Wang, S., Mieczkowski, J., Drier, Y., Gillespie, S.M., Archer, T.C., Wu, J.N., et al. (2017). SMARCB1-mediated SWI/SNF complex function is essential for enhancer regulation. *Nat Genet* 49, 289-295.
61. Raab, J.R., Runge, J.S., Spear, C.C., and Magnuson, T. (2017). Co-regulation of transcription by BRG1 and BRM, two mutually exclusive SWI/SNF ATPase subunits. *Epigenetics Chromatin* 10, 62.
62. Hodges, H.C., Stanton, B.Z., Cermakova, K., Chang, C.Y., Miller, E.L., Kirkland, J.G., Ku, W.L., Veverka, V., Zhao, K., and Crabtree, G.R. (2018). Dominant-negative SMARCA4 mutants alter the accessibility landscape of tissue-unrestricted enhancers. *Nat Struct Mol Biol* 25, 61-72.
63. Nguyen, H., Kerimoglu, C., Pirouz, M., Pham, L., Kiszka, K.A., Sokpor, G., Sakib, M.S., Rosenbusch, J., Teichmann, U., Seong, R.H., et al. (2018). Epigenetic Regulation by BAF Complexes Limits Neural Stem Cell Proliferation by Suppressing Wnt Signaling in Late Embryonic Development. *Stem Cell Reports* 10, 1734-1750.
64. Sokpor, G., Xie, Y., Rosenbusch, J., and Tuoc, T. (2017). Chromatin Remodeling BAF (SWI/SNF) Complexes in Neural Development and Disorders. *Front Mol Neurosci* 10, 243.
65. Son, E.Y., and Crabtree, G.R. (2014). The role of BAF (mSWI/SNF) complexes in mammalian neural development. *Am J Med Genet C Semin Med Genet* 166C, 333-349.
66. Kosho, T., Miyake, N., and Carey, J.C. (2014). Coffin-Siris syndrome and related disorders involving components of the BAF (mSWI/SNF) complex: historical review and recent advances using next generation sequencing. *Am J Med Genet C Semin Med Genet* 166C, 241-251.
67. Gazdag, G., Blyth, M., Scurr, I., Turnpenny, P.D., Mehta, S.G., Armstrong, R., McEntagart, M., Newbury-Ecob, R., Tobias, E.S., Study, D.D.D., et al. (2018). Extending the clinical and genetic spectrum of ARID2 related intellectual disability. A case series of 7 patients. *Eur J Med Genet*.

68. Zarate, Y.A., Bhoj, E., Kaylor, J., Li, D., Tsurusaki, Y., Miyake, N., Matsumoto, N., Phadke, S., Escobar, L., Irani, A., et al. (2016). SMARCE1, a rare cause of Coffin-Siris Syndrome: Clinical description of three additional cases. *Am J Med Genet A* 170, 1967-1973.
69. Khan, U., Study, D., Baker, E., and Clayton-Smith, J. (2018). Observation of Cleft Palate in an Individual with SOX11 Mutation: Indication of a Role for SOX11 in Human Palatogenesis. *Cleft Palate Craniofac J* 55, 456-461.
70. Wang, W., Xue, Y., Zhou, S., Kuo, A., Cairns, B.R., and Crabtree, G.R. (1996). Diversity and specialization of mammalian SWI/SNF complexes. *Genes Dev* 10, 2117-2130.
71. Kazantseva, A., Sepp, M., Kazantseva, J., Sadam, H., Pruunsild, P., Timmusk, T., Neuman, T., and Palm, K. (2009). N-terminally truncated BAF57 isoforms contribute to the diversity of SWI/SNF complexes in neurons. *J Neurochem* 109, 807-818.

| Individual # | # 1 | # 2 | # 3 | # 4* | # 5 | # 6 | # 7 |
|--|-------------|---|------------------|---------------------|---|---|---|
| gender | M | M | M | F | M | F | M |
| age (Y) | 5 Y | 3 Y | 2 Y | 17 Y | 4 Y | 18 Y | 11 Y |
| nucleotide change | c.400 A>G | c.723G>A | c.999dupA | c.1311-3C>G | c.1826T>C | c.1829T>C | c.1829T>C |
| amino acid change | p.Asn134Asp | p.Trp241* | p.Glu334Argfs*49 | splicing variant | p.Leu609Pro | p.Leu610Pro | p.Leu610Pro |
| de-novo variant? | yes | no (affected father) | n/a | yes | yes | yes | yes |
| affected domain | SMARCC_N | SMARCC_N | SMARCC_N | SWIRM | SANT | SANT | SANT |
| Clinical information | | | | | | | |
| developmental delay and/or intellectual disability | mild DD | - | mild DD | severe DD | severe DD | severe DD | severe DD |
| speech impairment | - | - | - | absence of language | absence of language | absence of language | absence of language |
| behavioral abnormalities | - | + | + | + | + | - | + |
| hypotonia | + | spasticity | + | + | + and spasticity | + | + |
| seizures | - | - | - | + | + | - | + |
| movement disorder | - | + | - | - | - | + | - |
| CNS abnormalities on MRI | n/a | two discrete hyperintens white matter lesions | n/a | normal MRI | thinning of corpus callosum and splenium, periventricular white matter loss | generalized cerebral atrophy, hypointensity globus pallidus | generalized cerebral atrophy, hypomyelination |
| Growth | | | | | | | |
| FTT | - | - | - | - | + | + | + |
| sucking/feeding difficulty | - | - | - | - | + | + | + |
| Craniofacial features | | | | | | | |
| thin/sparse scalp hair | - | - | - | - | + | + | - |

| | | | | | | | |
|---|-----|---|---|----------|--------|--------------------------|--|
| hypertrichosis | - | - | - | + | + | - | - |
| thick eyebrows | - | - | - | + | - | + | + |
| long eyelashes | - | - | - | - | - | - | - |
| ptosis | - | - | - | - | - | + | + |
| thin upper lip vermillion | - | - | - | - | + | + | - |
| thick lower lip vermillion | - | - | - | + | - | - | + |
| palate abnormalities | - | - | - | - | - | - | - |
| nose upturned/anteverted nostrils | - | - | - | - | + | + | + |
| Skeletal-limb features | | | | | | | |
| 5th finger or toe / nails abnormalities | - | - | - | - | + | - | - |
| scoliosis | - | - | - | kyphosis | + | + | + |
| Other | | | | | | | |
| cardiovascular | n/a | - | - | - | n/a | left coronary distension | non-progressive mild aortic dilatation (Z score=2.54). |
| inguinal hernia | - | - | - | - | + | - | + |
| undescended testis | n/a | - | - | n/ap | + | n/ap | + |
| skin problems | - | hypo pigmented hair, cafe au lait macules | - | - | eczema | eczema, scleroderma | - |

| # 8 | # 9 | # 10 ** | # 11 | # 12 | # 13 | # 14 | # 15 | Total |
|---------------------------------------|------------------------------------|--|--------------------------------|---|--------------------------|--------------------------|-----------------------------------|-------|
| M | M | M | M | F | M | M | M | |
| 11 Y | 2.5 Y | 10.5 Y | 19 Y | 7 Y | 8 Y | 5 Y | 27 Mo | |
| c.1833+1G>T splicing variant*** | c.1833+2T>C splicing variant | c.1833+2T>C splicing variant | c.1838T>C p.Leu613Pro | c.1903A>G p.Cys635Arg | c.2686A>G p.Met896Val | c.2699A>G p.Glu900Gly | c.3456_3458delCAT p.Met1153del | |
| n/a | yes | yes | yes | yes | yes | yes | yes | |
| SANT | SANT | SANT | SANT | SANT | SMARCC_C | SMARCC_C | carboxy terminal | |
| severe DD (DQ=20) | moderate ID | moderate DD, moderate- severe ID | moderate- severe DD | moderate DD, moderate ID | moderate- severe DD | mild ID | mild DD | 14/15 |
| absence of language | absence of language | minimal speech | absence of language | speech delay | speech delay | speech delay | speech abnormalities | 12/15 |
| + | - | - | - | + | + | + | + | 10/15 |
| + | + | + | - | + | + | + | + | 13/15 |
| - | - | - | + | - | - | - | - | 4/15 |
| - | - | - | - | - | - | - | - | 2/15 |
| normal MRI | normal MRI | normal MRI | abnormal corpus callosum | slightly small corpus callosum, prominent perivascular spaces | normal MRI | normal MRI | n/a | 6/12 |
| - | + | + | - | - | - | + | - | 6/15 |
| - | + | + | + | - | - | + | + | 8/15 |
| | | | | | | | | 11/15 |
| - | + | - | + | - | - | - | - | 4/15 |

| | | | | | | | | |
|---|-------------------------|--|---|--------|-----|---|-----|------|
| - | + | - | - | + | + | + | - | 6/15 |
| - | + | - | + | - | - | + | - | 6/15 |
| - | + | + | + | - | + | + | - | 5/15 |
| - | - | + | + | - | - | + | - | 5/15 |
| + | + | - | - | + | + | - | - | 6/15 |
| - | + | - | + | - | + | - | - | 5/15 |
| - | - | + | + | - | - | + | - | 3/15 |
| - | - | + | + | + | - | - | - | 6/15 |
| | | | | | | | | |
| - | - | + | + | - | - | + | - | 4/15 |
| - | - | - | - | + | - | - | - | 5/15 |
| | | | | | | | | |
| - | - | - | - | - | n/a | - | n/a | 2/11 |
| + | - | - | - | - | - | - | - | 3/15 |
| - | n/a | - | - | n/ap | - | - | n/a | 2/12 |
| - | hypomelanotic macula | vitiligo (present in unaffected father) | hyper pigmented irregular skin on back | eczema | - | - | - | 7/15 |

Table 1: Summary of *SMARCC2* variants and clinical presentation of 15 individuals. * see Martinez et. al; ** see Zhu et.al; *** see figure 3; '-' - not reported in this individual; '+' - reported in this individual; FTT – failure to thrive; DD – developmental delay; ID – Intellectual disability; DQ - Developmental Quotient; CNS – central nervous system; MRI – magnetic resonance imaging; M - male; F - female; Y - years; Mo - months; n/a - not available information; n/ap - not applicable.

| | | ARID1A | ARID1B | ARID2 | SMARCA2 | SMARCA4 | SMARCB1 |
|--|---------------------------------------|--|--------------------|--------------------|--------------------|--|--|
| Associated Condition (MIM#) | | CSS 2 (614607) | CSS 1 (135900) | CSS 6 (617808) | NCBRS (601358) | CSS 4 (614609) | CSS 3 (614608) |
| OMIM Gene (*) | | 603024 | 614556 | 609539 | 600014 | 603254 | 601607 |
| References (PMID) | | Koshko et al, 2014; Koshko, Miyake, and Carey 2014; Santen et al, 2013 | Santen et al, 2013 | Gazdag et al, 2018 | Santen et al, 2013 | Koshko et al, 2014; Koshko, Miyake, and Carey 2014; Santen et al, 2013 | Koshko et al, 2014; Koshko, Miyake, and Carey 2014 |
| Types of pathogenic mutations described in literature | missense | | | | 4/4 | 10/12 | 4/13 |
| | in-frame deletion | | | | | 2/12 | 9/13 (all p.Lys364del) |
| | splicing mutation | | 2/29 | | | | |
| | frameshift and stop gained / nonsense | 8/8 | 25/29 | 11/14 | | | |
| | partial/full gene deletions | | 2/29 | 3/14 | | | |
| Short stature (<= -2 SD) | | 3/8 | 3/27 | 10/14 | 1/4 | 9/12 | 11/11 |
| Weight <=-2 SD | | 3/6 | 11/28 | | 2/4 | 4/12 | 8/9 |
| OFC < -2SD | | 2/8 | 0/27 | 0/14 | 1/4 | 10/11 | 9/10 |

| | | | | | | | |
|--|------------------------|-------------|-------|---|-----|--------------------------|-----------------|
| DD and/or ID | | 7/7 | 28/28 | 14/14 | 4/4 | 11/11 | 11/11 |
| Speech delay (first word after 12 months) | | 6/6 | 28/28 | 11/11 | 4/4 | 11/11 | 10/10 |
| Severe speech delay (1st word =>5)/No language | | 5/6 | 17/28 | 5/11 | 2/4 | 6/11 | 9/10 |
| Seizures | | 2/7 | 5/20 | 3/4 | 0/4 | 2/12 | 8/10 |
| Hypotonia | | 7/8 | 23/27 | 2/14 | 2/4 | 8/11 | 8/11 |
| Brain MRI abnormalities | | 7/8 | 9/25 | 6/6 | 0/3 | 6/7 | 9/9 |
| Behavioral phenotypes | | 3/5 | n/r | 10/14 | n/r | 7/8 | 4/8 |
| | phenotype described | AX, OCB, HA | n/r | ADHD, AG, AX, HA, HS (to loud noises), RG, SLD, STM | n/r | HA, HS, IM, OCB, SAS, SI | HA, IM, SI, TAN |
| Dysmorphic features | reported as coarse | n/r | n/r | 9/14 | n/r | n/r | n/r |
| | thick eyebrows | 6/8 | 24/27 | 2/7 | 4/4 | 9/12 | 11/11 |
| | thick lower vermillion | 6/7 | 22/24 | 1/7 | 2/3 | 10/12 | 8/11 |
| | hypertrichosis | 7/7 | 26/28 | 0/7 | 4/4 | 12/12 | 8/11 |
| | sparse scalp hair | 3/5 | 18/28 | 0/7 | 4/4 | 5/12 | 10/11 |
| Cleft palate | | 2/6 | n/r | 1/14 | n/r | 5/12 | 2/13 |
| Laryngotracheal anomalies | | 1/4 | 2/28 | 2/14 | n/r | 2/4 | n/r |
| Hearing loss | | 2/6 | 5/22 | n/r | 0/3 | 4/12 | 8/11 |

| | | | | | | | |
|---|---|-----|-------|------|-----|-------|-------|
| Hand, foot & digital anomalies | hypoplasia or absence of 5th distal phalanx | 6/7 | 5/22 | 2/10 | 0/3 | 10/12 | 8/11 |
| | Hypoplastic nails | 6/8 | 19/28 | 8/12 | 0/2 | 10/12 | 11/11 |
| | Brachydactyly | 0/8 | 11/24 | 1/10 | 1/3 | n/r | n/r |
| | Clinodactyly | n/r | n/r | 0/7 | n/r | n/r | n/r |
| | Syndactyly | n/r | n/r | 0/7 | n/r | n/r | n/r |
| Other skeletal anomalies (including scoliosis) | | 2/7 | n/r | 4/14 | n/r | 1/10 | 7/9 |
| Cryptorchidism | | 2/5 | n/r | n/r | n/r | 3/9 | 2/5 |
| Heart defect | | 3/8 | n/r | 1/14 | | 5/12 | 5/11 |
| | dextrocardia | 0/8 | 0/27 | n/r | 0/4 | 0/12 | 2/11 |

| SMARCE1 | SOX11 | ACTL6A | DPF2 | SMARCC2 |
|--|---|----------------------|-------------------------|-------------------|
| CSS 5 (616938) | Mental retardation, AD 27 (615866) | ID syndrome | CSS 7 (618027) | CSS -like |
| 603111 | 600898 | 604958 | 601671 | 601734 |
| Koshko et al, 2014; Kosho, Miyake, and Carey 2014; Zarate et al, 2016 | Khan et al, 2018 | Marom et al, 2017 | Vasiliou et al, 2018 | current report |
| 6/6 | 4/13 | 2/3 | 5/8 | 8/15 |
| | | | | 1/15 |
| | | 1/3 | 1/8 | 4/15 |
| | 2/13 | | 2/8 | 2/15 |
| | 7/13 | | | |
| 3/5 | 4/11 | 0/3 | 4/8 | 6/13 |
| 4/6 | n/r | 0/3 | 4/7 | 5/13 |
| 4/6 | 3/13 | n/r | 0/7 | 6/13 |
| 6/6 | 10/13 | 3/3 | 8/8 | 14/15 |

| | | | | |
|------------|------------------------------------|------------------------------------|---|--|
| 2/6 | 11/13 | 3/3 | 8/8 | 12/15 |
| n/r | n/r | 0/3 | n/r | 8/15 |
| 2/6 | 4/14 | 0/3 | n/r | 4/15 |
| 2/6 | 3/13 | 0/3 | 4/7 | 13/15 |
| 3/4 | 2/3 | n/r | 3/3 | 6/12 |
| 1/1 | 6/13 | 2/3 | 3/8 | 10/15 |
| HA | ASD, ADHD,AG, HA, SI, SLD | ADD/ADHD, AG, IM,TAN, SLD | STM, TAN, OCB, HA, SLD, RG, DSI/ASD | ADHD, AG, ASD, AX, DSI, HA, HP, HS, OCB, RG, SI , SLD, TAN |
| 6/6 | n/r | 1/3 | 2/8 | n/r |
| 5/6 | 1/13 | 1/3 | n/r | 6/15 |
| 6/6 | 7/13 | n/r | n/r | 5/15 |
| n/r | 2/13 | n/r | n/r | 6/15 |
| 5/6 | 4/13 | n/r | 6/7 | 4/15 |
| 2/6 | 1/13 | 0/3 | n/r | 0/15 |
| 1/6 | n/r | 1/3 | n/r | 2/15 |
| 2/6 | 2/14 | n/r | 4/8 | 2/15 |
| 2/6 | n/r | 0/3 | 0/8 | 0/15 |
| 6/6 | 7/14 | 1/3 | 8/8 | 3/15 |
| 0/6 | 0/14 | 1/3 | 5/8 | 0/15 |
| 0/6 | 12/14 | 2/3 | 3/8 | 1/15 |

| | | | | |
|------------|------|-----|---------------------------|------|
| 0/6 | 4/14 | 1/3 | n/r | 0/15 |
| 3/6 | 4/14 | 1/3 | 3/8 (craniosynostosis) | 7/15 |
| 1/4 | 3/4 | 0/3 | n/r | 2/12 |
| 4/5 | | 1/3 | 4/8 | 2/11 |
| 1/5 | n/r | 0/3 | n/r | 0/11 |

Table 2: Summary of molecular variants and clinical presentation in genes associated with Coffin-Siris and Coffin-Siris-like syndromes.

AD - Autosomal Dominant; ADHD - attention deficit hyperactivity disorder; AG - aggressiveness; ASD - Autism Spectrum Disorder; AX - anxiety; CCS - Coffin-Siris syndrome; DD - developmental delay; DSI-difficulties with social interactions (not qualify for ASD diagnosis); HA - hyperactivity; HP - Hyperphagia; HS - hypersensitive; ID - Intellectual disability; IM - impulsive; MRI - magnetic resonance imaging; NCBRS - Nicolaides-Baraitser syndrome; n/r - not reported; OCB - obsessive-compulsive behavior; OFC - occipital frontal circumference; RG - rigidity/‘routine driven’; SAS - short attention span; SD - standard deviation; SI - self-injury; SLD - sleep disturbance; STM - stereotypic movements; TAN - Tantrums; For comparison of clinical presentation of other BAF related genes (including *PHF6*, *ADNP*, and *TBC1D24*) see Koshko et al, 2014 ⁶⁶.

3.3 Des mutations dans *CHD3* sont associées à un syndrome neurodéveloppemental et induisent des changements d'expression génique

Nous avons identifié 35 individus avec 27 mutations différentes dans le gène *CHD3* qui présentaient un syndrome neurodéveloppemental. Nous avons introduit par mutagenèse dirigée six des mutations les plus récurrentes dans un plasmide d'expression avec une étiquette eGFP. Nos collaborateurs dans le laboratoire Wade au NIEHS ont exprimé puis purifié ces constructions, et démontré par un essai d'activité ATPase et un essai de déplacement de nucléosomes que sur ces six mutations, quatre ont une activité de remodelage de nucléosomes défectueuse, et une a une activité augmentée²⁰⁰. De ces mutations à activité différente de la protéine sauvage, trois ont une activité ATPase amoindrie tandis que la mutation donnant une activité de remodelage de nucléosomes augmentée est aussi associée à une activité ATPase plus importante. Ainsi, ils ont démontré que ces mutations induisent des défauts dans l'activité de *CHD3*.

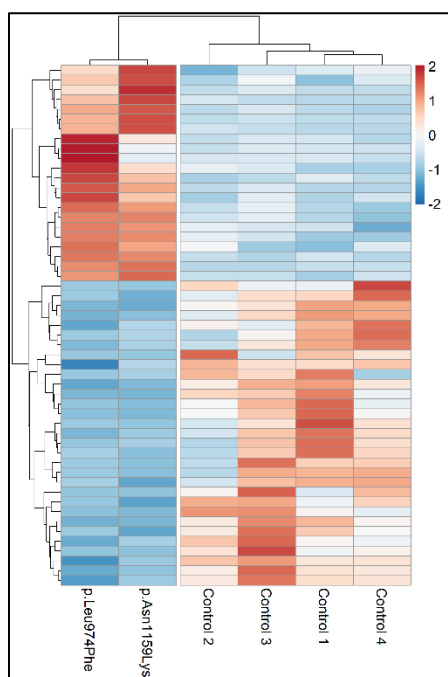


Figure 16: Les fibroblastes porteurs de mutations dans *CHD3* présentent un profil d'expression génique différent. Des fibroblastes provenant de deux individus porteurs respectivement des mutations NM_001005273.2 : p.L915F et p.N1159K ont été établis, et leur ARN total a été séquençé par RNAseq sur une librairie enrichie en queues polyA,

avec un échantillon par lignée. L'analyse bioinformatique est détaillée dans la section des méthodes. Les 59 gènes différentiellement exprimés par les deux fibroblastes porteurs de mutations dans *CHD3* par rapport à quatre contrôles ($\text{padj} < 0.01$, $\text{Log2FC} > 2$) sont visualisés par une heatmap où l'expression des gènes plus élevée est représentée par la couleur rouge, tandis que l'expression moins élevée est représentée par la couleur bleue.

Nous avons pu obtenir des fibroblastes provenant de deux individus porteurs de ces mutations, soit la mutation p.L915F, causant une activité augmentée, et p.N1159K, ayant une activité presque inexistante. Nous nous sommes donc demandé si des défauts dans le remodelage de la chromatine causeraient des différences d'expression génique et de la structure de la chromatine. L'activité augmentée ou réduite du déplacement des nucléosomes le long de l'ADN pourrait causer des changements de l'expression génique similaires, où les loci ciblés auraient une structure anormale. Nous avons donc analysé l'expression des deux lignées en les considérant comme répliquats de la même condition, ce qui correspond à 59 gènes différentiellement régulés dans les deux individus ($\text{log2FC} > 2$, $\text{padj} < 0.01$, Figure 16). En faisant une revue de la littérature de ces gènes, on observe que presque la moitié sont associés au développement neural ou à la fonction neuronale (Tableau VI).

Tableau VI - Gènes différentiellement régulés dans des fibroblastes d'individus porteurs de mutations dans *CHD3* impliqués dans le développement neural. Les 59 gènes différentiellement exprimés dans les fibroblastes provenant d'individus porteurs d'une mutation dans *CHD3* ont été associés à une fonction par revue manuelle de la littérature. 27 gènes ont été sélectionnés pour leur fonction dans le développement neural, la fonction neuronale, ou le cytosquelette/adhérence cellulaire.

| Gène | log2FoldChange | padj | Fonction |
|---------|----------------|----------|------------------------------------|
| AJAP1 | 5.749532 | 1.72E-14 | Cytosquelette/Adhérence Cellulaire |
| RIMS1 | -5.53026 | 9.02E-09 | Fonction Neuronale |
| IGSF10 | 3.901759 | 7.05E-08 | Développement Neural |
| PPP2R2B | -5.7364 | 2.95E-06 | Développement Neural |
| SCRG1 | 4.471867 | 3.72E-06 | Développement Neural |
| SPON1 | 3.305276 | 2.49E-05 | Fonction Neuronale |
| PDE11A | -4.98013 | 4.59E-05 | Fonction Neuronale |
| EDIL3 | -4.21355 | 5.32E-05 | Développement Neural |
| ABLIM1 | -2.74679 | 0.000115 | Fonction Neuronale |
| OPCML | -3.51261 | 0.000171 | Développement Neural |
| SYT1 | -2.88565 | 0.000185 | Fonction Neuronale |
| TRH | 4.662962 | 0.000228 | Fonction Neuronale |
| PPL | 3.047235 | 0.000448 | Fonction Neuronale |

| | | | |
|-----------------|----------|----------|------------------------------------|
| SLC6A17 | -3.22161 | 0.000566 | Fonction Neuronale |
| GRIA4 | 4.744157 | 0.000789 | Développement Neural |
| PPP1R14C | -4.35755 | 0.001234 | Fonction Neuronale |
| ST8SIA5 | -3.27798 | 0.001256 | Fonction Neuronale |
| TMEM150C | -3.36194 | 0.001435 | Développement Neural |
| ALS2CL | -2.37697 | 0.001735 | Fonction Neuronale |
| PRUNE2 | -2.68799 | 0.002449 | Développement Neural |
| MCF2L | -3.7851 | 0.002604 | Fonction Neuronale |
| CELSR1 | -3.49912 | 0.002805 | Développement Neural |
| SAMD12 | -2.09953 | 0.003035 | Fonction Neuronale |
| PARD6G | -2.29205 | 0.003532 | Cytosquelette/Adhérence Cellulaire |
| NTNG1 | -2.66482 | 0.005345 | Développement Neural |
| PPP1R9A | -3.91374 | 0.005345 | Développement Neural |
| SLC1A1 | -3.2517 | 0.009457 | Fonction Neuronale |

3.4 Analyse de l'ouverture de la chromatine dans des fibroblastes d'individus *TRRAP*, *SMARCC2*, et *CHD3*

Ayant démontré que les mutations dans les gènes *TRRAP*, *SMARCC2*, et *CHD3* causaient des changements d'expression dans les cellules d'individus affectés comparés à des contrôles sains, on a voulu étudier leur impact au niveau de la structure de la chromatine. Pour ce faire, nous sommes partis de l'hypothèse que des défauts dans le remodelage de la chromatine ou dans l'acétylation d'histones pourraient engendrer des différences dans les régions d'ouverture de la chromatine. Ces différences pourraient exister dans deux conditions : premièrement, le défaut de remodelage de nucléosomes pourrait induire des changements dans la localisation des régions d'ouverture de la chromatine où les pics correspondant à des régions libres de nucléosomes seraient décalés dans les cellules d'individus affectés. Aussi, ces défauts pourraient induire des changements dans la structure de la chromatine, où les pics correspondant à la chromatine ouverte seraient de taille et d'intensité différente. Alors, nous avons décidé d'utiliser la méthode du séquençage ATAC (assay for transposase accessible chromatin) qui nous permettrait d'identifier les régions ouvertes de la chromatine et de les comparer à des cellules provenant d'individus sains.

Nous avons décidé de comparer les mêmes cellules séquencées par RNAseq afin d'établir une corrélation entre DEGs et ouverture de la chromatine. Ainsi, nous avons procédé à l'analyse par ATACseq des fibroblastes provenant d'individus porteurs de mutations dans *CHD3*, *TRRAP*, ou *SMARCC2*. Comme la transposase favorise l'ADN ouvert et que le chromosome mitochondrial n'est pas condensé, nous avons enlevé les séquences correspondant à l'ADN mitochondrial, qui correspondaient à 40-60% des fragments séquencés. Après les étapes de nettoyage des fichiers, il restait environ 5 à 15 millions de fragments, correspondant à 90 000-120 000 pics larges.

L'analyse subséquente a été réalisée avec le logiciel Homer, qui a permis de confirmer qu'il y avait un enrichissement de pics dans les régions correspondant aux ARN (exon, intron), promoteurs, et 5'UTR, ce qui correspond aux attentes où les régions ouvertes de la chromatine sont celles où la machinerie de transcription a accès. Cependant, afin de faire une analyse statistique, il est nécessaire de considérer les trois contrôles et les individus porteurs de mutations dans le même gène comme des répliquats des mêmes conditions, ce que le logiciel détectait comme étant problématique, où l'auto-corrélation entre les échantillons était faible et qu'il y avait un enrichissement amoindri. Alors, l'analyse statistique n'a pu détecter qu'entre 0 et 300 pics différents entre contrôles et individus affectés, où la quasi-totalité des pics différents se trouve dans des introns ou dans des régions intergéniques. De plus, en considérant une annotation des pics au gène le plus proche, correspondant probablement à des régions de promoteur ou de enhancer (région connue comme ayant un impact positif sur la transcription d'un gène ou d'un groupe de gènes), très peu des pics se rapprochent de gènes différentiellement exprimés.

Nous avons procédé à l'analyse pour l'ATACseq pour seulement une des deux lignées de fibroblastes d'individus TRRAP (p.W1866C, Annexe II). Il est intéressant que l'on trouve onze gènes qui sont différentiellement exprimés dans les deux fibroblastes TRRAP, et qu'il existe des pics significativement plus grands au niveau de leur promoteur dans la lignée séquencée par ATACseq (Tableau VII). De ces gènes, presque la moitié est impliquée dans le neurodéveloppement ou le fonctionnement neuronal, notamment *SP9*, *DRD1*, *TRIM36*, *DYNC1H1*, et *FGF12*.

Tableau VII - Gènes significatifs en RNAseq et ATACseq pour des fibroblastes TRRAP. Les 285 gènes significatifs en RNAseq ($\text{Log2FC} > 2$ et $\text{padj} < 0.01$) ont été comparés 359 aux pics significatifs en ATACseq ($\text{padj} < 0.05$) annotés au gène ayant le TSS le plus proche.

| Gène | log2FoldChange RNAseq | padj RNAseq | Log2FoldChange ATACseq | padj ATACseq |
|-----------------|-----------------------|-------------|------------------------|--------------|
| SP9 | 4.93073586 | 5.93E-11 | 2.74037472 | 0.03587464 |
| PAX9 | 2.39854929 | 8.22E-09 | 3.42667259 | 0.02100832 |
| PARP8 | 5.17592124 | 1.34E-08 | 2.16295983 | 0.03557351 |
| EMB | 8.38520896 | 1.90E-06 | 3.40159632 | 0.04352787 |
| DRD1 | 3.42449606 | 0.00035459 | 2.7035415 | 0.02016334 |
| APBB1IP | 2.60235882 | 0.00047851 | 2.97705712 | 0.04688938 |
| GATA6 | 2.5823747 | 0.00079646 | 4.0152643 | 0.04940477 |
| FGF12 | 3.18762447 | 0.00359546 | 3.7181421 | 0.04196407 |
| KIAA1671 | 2.54056845 | 0.00403068 | 6.91262031 | 0.02728563 |
| TRIM36 | 2.57511414 | 0.0056605 | 2.86346218 | 0.04940477 |
| DYNC1I1 | 2.60108211 | 0.00585867 | 5.32955024 | 0.0463569 |

Alors, il est possible que les défauts dans l'expression génique des individus TRRAP soient causés par une ouverture anormale de la chromatine au niveau du promoteur de ces gènes mal régulés. Pour les autres individus (SMARCC2 et CHD3), nous n'avons pas trouvé de corrélation entre RNAseq et ATACseq.

4. Discussion

Sachant que le développement neural est un concours de mécanismes extrêmement précis dépendant de changements d'expression génique subtils, de nombreuses mutations dans des régulateurs de la chromatine ont été associées à des syndromes neurodéveloppementaux. Nous avons caractérisé trois syndromes associés aux régulateurs de la chromatine *TRRAP*, *SMARCC2*, et *CHD3*. Pour chacun, nous avons démontré que dans des cellules provenant de ces individus (fibroblastes ou LCLs), un profil de gènes dérégulés émergeait et que ceux-ci sont en grande partie impliqués dans le développement neural ou le fonctionnement neuronal. Bien qu'elles soient des protéines exprimées ubiquitairement, il est probable que leur impact est majoritairement sur le développement neural dû à sa nécessité d'un système fin et précis où chaque cellule est nécessaire pour le fonctionnement normal du cerveau. Dans d'autres systèmes, il y a peut-être une plus grande tolérance à ces défauts d'expression, et leur fonctionnement ne nécessiterait pas la totalité des cellules fonctionnelles pour être opérationnel. Enfin, les gènes affectés qui causent un phénotype cellulaire pourraient nécessiter la combinaison avec des facteurs de transcription qui ne sont exprimés que lors du développement neural.

4.1 Limitations techniques et améliorations

Un important défaut de nos études est la taille de l'échantillon particulièrement petit. En effet, il est particulièrement difficile d'obtenir le consentement des parents des individus affectés afin de procéder à une biopsie de la peau, relativement invasive, ou même à obtenir du sang. De ce fait, bien qu'on ait caractérisé entre 15 et 35 individus pour chaque syndrome, seuls deux à trois ont accepté de donner des tissus. Alors, les résultats que nous obtenons sont extrêmement biaisés. En ayant séquencé le même individu à deux moments différents, on a donc changé plusieurs paramètres, soit le passage des cellules (différence de 3 passages), les contrôles auxquels il était comparé (deux contrôles versus quatre contrôles différents), le kit de préparation de librairie (Illumina TruseqV2 vs NEBNext Ultra II directional), et la méthode de séquençage (Illumina HiSeq4000 vs NovaSeq S2). Les deux derniers paramètres devraient affecter seulement un minimum le séquençage, et ne devraient compter que pour une infime partie des différences. Cependant, en comparant des passages différents, et à des contrôles

différents, on change complètement le cadre de l'analyse statistique, et les résultats sont extrêmement différents. En comparant seulement le fibroblaste répété (TRRAP p.L805F) avec les contrôles des expériences respectives, on obtient 343 et 270 DEGs, dont seulement 90 sont communs.

Ces changements sont vérifiables en augmentant le nombre de contrôles pour obtenir une vision plus proche de la moyenne de l'expression génique dans des cellules contrôles, et en étudiant les cellules à passages différents. En effet, c'est un exercice qui a été fait par qPCR sur plusieurs gènes (Figure 10 et résultats non-montrés), et qui démontre qu'il y a un biais quant aux contrôles utilisés, où les résultats ne sont pas tous reproductibles avec plus de contrôles, et où l'écart-type de la valeur d'un gène dans la même lignée est relativement élevé. On a pu observer que pour la première expérience de séquençage où seulement deux contrôles différents étaient utilisés, plus de 50% des gènes qui faisaient partie des gènes les plus significativement dérégulés n'étaient plus significativement dérégulés en rajoutant des contrôles et des réplicats (six gènes testés dans deux individus). Dans notre deuxième expérience de séquençage, passer de deux contrôles à quatre a permis de réduire cette proportion à 20% (cinq gènes testés sur deux individus SMARCC2).

Les données d'ATACseq n'ont été que peu conclusive, particulièrement à cause de la variabilité inter-individuels. Il est important de noter que les fibroblastes provenant de biopsies différentes ont des caractéristiques particulières, et expriment des marqueurs uniques²⁵². Ceci suggérerait un profil d'expression singulier à chaque sous-type de fibroblaste et donc une ouverture de la chromatine différente. Or, les fibroblastes utilisés proviennent généralement de deux tissus relativement différents (fibroblastes de la peau et fibroblastes du prépuce). En refaisant l'expérience avec plus de réplicats par individu, et des contrôles provenant du même tissu que les individus étudiés, on pourrait peut-être procéder à des comparaisons plus représentatives du tissu source, et se rapprochant des vraies valeurs. Alors, on pourrait faire des analyses plus poussées en s'intéressant à des régions spécifiques du génome. Par exemple, on pourrait comparer les pics d'ouverture de la chromatine à des données de méthylation H3K4me1/me3 qui caractérisent les séquences des promoteurs (haut taux de H3K4me3 / bas taux de H3K4me1) et des enhancers (haut taux de H3K4me1 / bas taux de H3K4me3)²⁵³.

Aussi, il est compliqué de déterminer un profil d'expression qui serait affecté par toutes les mutations de la même façon. Sachant que les protéines qu'on étudie sont présentes dans plusieurs complexes de remodelage de la chromatine, et qu'ils interagissent avec une multitude de protéines, il est possible que chaque mutation ait un impact différent. Alors, le seul moyen de les étudier serait dans un contexte isogénique, soit de procéder à une ablation génétique du gène, puis de réintroduire ses versions sauvages et mutantes. Cependant, il faut aussi faire attention à ce que l'expression du gène soit à un niveau similaire à celui de la protéine endogène, particulièrement dans le cas de SMARCC2 qui agit en compétition avec SMARCC1, et où la quantité relative de chacune des protéines est variable au cours du développement, démontrant une nécessité de garder des niveaux d'expression distincts. Des expériences ont déjà été réalisées *in vitro* pour démontrer les différences d'activité hélicase de CHD3, où différentes mutations résultent en des taux d'activité spécifiques, et même une activité plus grande ou amoindrie selon la localisation de la mutation.

Enfin, la comparaison entre la même mutation dans deux lignées cellulaires démontre des effets dépendants du type cellulaire. Effectivement, des LCLs et des fibroblastes provenant du même individu ne montrent pas de DEGs en commun. De plus, on n'a pu démontrer un profil spécifique au complexe BAF en comparant tous les sets de données sur le GEO omnibus ayant rapport avec l'une de ses sous-unités¹⁹⁰. Alors, il est difficile de confirmer les résultats qu'on a obtenu. Ainsi, pour avoir des résultats significatifs par rapport aux syndromes étudiés, on a décidé de dé-différencier les cellules provenant d'individus en cellules pluripotentes induites (iPSCs) et de les différencier en CSNs afin d'étudier un modèle plus approprié.

4.2 Contexte et signification

L'étude de lignées primaires ou transformées de patients porteurs de mutations dans des régulateurs de la chromatine associées à des maladies neurodéveloppementales n'a jamais été faite auparavant. Ainsi, il est difficile d'évaluer nos résultats. Il a été démontré récemment que des profils de méthylation spécifiques étaient retrouvés dans l'ADN provenant du sang d'individus affectés de TDISS, et que les régions différentiellement méthylées de ceux-ci étaient associées notamment au développement embryonnaire, à la régulation des processus

développementaux, et à la maturation synaptique²⁵⁴. Comme pour nos études, bien qu'étant dans des cellules non-spécifiques au cerveau et établies après le développement embryonnaire, ces gènes et leurs régions régulatrices sont particulièrement enrichis dans les profils des individus affectés comparés aux contrôles sains. Grâce à un modèle spécifique de méthylation, les auteurs ont pu caractériser la pathogénicité d'un variant précédemment identifié comme « non-concluant » (Variant of Unknown Significance), en montrant que le profil de l'individu concerné était différentiellement méthylé comme ceux des autres individus affectés. En séquençant l'ARN de plus de patients affectés, nous espérons pouvoir établir un modèle similaire et aider au diagnostic d'individus présentant des variants dont le potentiel pathogénique est flou.

En général, nous nous sommes intéressés aux gènes impliqués dans la neurogenèse ou la fonction neuronale. Bien que les phénotypes observés chez les porteurs de variants dans les trois gènes étudiés sont à la fois très hétérogènes dans une même cohorte, mais ont plusieurs éléments en commun (DI, DD, TSA), il n'y a pas ou peu de gènes communs qui sont retrouvés significativement différentiellement régulés. Il semble qu'il s'agirait plutôt de groupes de gènes impliqués dans des voies de signalisation communes qui seraient affectés, notamment des gènes importants pour la formation du cytosquelette et l'adhésion cellulaire. Certains gènes des voies développementales sont dérégulés, tels les BMPs *GDF6* et *GDF10* dans les fibroblastes *SMARCC2*, ou les facteurs de la voie Hedgehog *NSG1* et *HHIP* dans les fibroblastes *TRRAP* (Annexe I). En revanche, aucune voie ne ressort significativement dérégulée dans un syndrome particulier. Alors, il est probable qu'une déviation du modèle normal du développement cortical avec un changement d'expression incluant une multitude de gènes serait suffisante pour causer un défaut pathogénique.

De façon intéressante, un gène qui est hautement dérégulé dans les cellules d'individus porteurs de mutations dans *TRRAP* et *SMARCC2* est *H19*. Ce long ARN non-codant a une empreinte réciproque au gène *IGF2*, qui est dépendante de la méthylation de leurs loci respectifs ainsi que de l'hypoacétylation de la région codant pour *H19*. *H19* régule la croissance cellulaire à travers des micro-ARNs : le 1^{er} exon de *H19* encode miR-675, et *H19* séquestre la famille de micro-ARNs let-7²⁵⁵⁻²⁵⁷. Pourtant, *IGF2*, la cible de miR-675 *IGF1R*, et les cibles de *Let-7* *DICER* et *HMGA2* ne sont pas significativement différentiellement régulés dans les cellules

d'individus porteurs des mutations dans *SMARCC2*, et *TRRAP*. Il est possible que d'autres cibles de ces miRNAs ou de l'ARN lui-même soient différentiellement régulées et se trouvent dans nos listes de DEGs.

Bien que nous nous soyons concentrés sur un défaut principalement neuronal, nous nous sommes rendu compte qu'une grande partie des DEGs trouvés dans les fibroblastes d'individus *TRRAP* et *SMARCC2* étaient impliqués dans le fonctionnement astrocytaire ou oligodendrocytaire (*SCRG1*, *PCDH7*, *SBTBN2*, ... voir Annexe). Il serait donc intéressant d'étudier non-seulement un modèle *in vitro* neuronal mais aussi des co-cultures avec des cellules gliales et dans un contexte *in vivo*, les gènes et les types cellulaires impliqués dans les pathogénèses. En effet, certaines maladies neurodéveloppementales impliquent des types cellulaires différents, comme la maladie de Pelizaeus-Merzbacher (OMIM : 312080) où l'hypomyélination par les oligodendrocytes cause un phénotype caractérisé par un retard du développement, l'hypotonie, et un nystagmus (mouvements répétitifs involontaires des yeux)²⁵⁸. De même, les variants de *FMRI* associés au syndrome de l'X fragile (OMIM : 300624) causent une dérégulation du transport du glutamate entre neurones et astrocytes par une dérégulation de l'expression astrocytaire du récepteur au glutamate *mGluR5* et du transporteur *EAAT-1*²⁵⁹⁻²⁶¹.

L'ARN des fibroblastes d'individus porteurs de variants dans *TRRAP* a été préparé selon une librairie différente des autres lignées étudiées. Cette librairie a permis non seulement le séquençage de l'ARNm, mais aussi d'autres ARNs porteurs d'une queue poly-A. Ainsi, dans les gènes différentiellement régulés se trouvent 227 gènes codants pour des protéines, mais aussi des longs ARNs non-codants (LncRNA, 6 gènes), des longs ARNs intergéniques non-codants (LincRNA, 12 gènes), des pseudogènes (20 gènes), et enfin des ARNs anti-sens (15 gènes). Nous nous sommes seulement intéressés aux gènes codants par manque de temps et d'annotations sur les gènes non-codants mais il est possible que leur dérégulation fasse partie du mécanisme pathogénique de la maladie. En effet, plusieurs LncRNAs ont été impliqués dans la progression de maladies neurodégénératives et dans le développement cortical^{262,263}.

4.3 Pistes Thérapeutiques

En identifiant les cibles des sous-unités de complexes épigénétiques, nous espérons trouver des gènes sous-jacents aux phénotypes observés chez les individus affectés. Des variants de *MECP2* (Methyl CpG Binding Protein 2) sont la majeure cause du syndrome de Rett (OMIM : 312750)²⁶⁴. Dans ce cas là, la régulation transcriptionnelle par *MECP2* et ses partenaires d'interaction est entravée et l'expression de la sous-unité du récepteur du glutamate *GRID1* est sous-exprimée²⁶⁵. Cette expression trop basse induit une malformation des synapses activatrices glutamatergique et une baisse des courants synaptiques de neurones *in vitro*. Le traitement avec l'acide valproïque, un activateur de la déméthylation de l'ADN et un inhibiteur global de la déacétylation, permet de rétablir la fonction et la morphologie de neurones *in vitro*, ainsi que de réduire les symptômes de TSA dans un modèle murin et de rétablir un taux d'expression génique normal pour une grande partie des gènes affectés dans le cerveau²⁶⁴.

De même, des essais cliniques sont en cours pour le traitement du syndrome de Rubinstein-Taybi, qui est causé par des variants dans l'histone acétyltransférase *CREBBP*, avec des inhibiteurs d'acétylation (Essai RUBIVAL ClinicalTrials ID : NCT01619644). Aucuns résultats n'ont été publiés pour l'instant, mais il est fort probable qu'un traitement serait efficace. En effet, il a aussi été montré qu'une hypoacétylation globale du cortex préfrontal murin causée par l'inactivation hétérozygote de *SHANK3*, un gène muté dans presque 2% des cas de TSA, est rétablie lors du traitement avec l'inhibiteur d'histone déacétylases Romidepsin²⁶⁶. La récupération de taux normaux d'acétylation est associée avec une diminution des symptômes associés au TSA chez ces souris. Alors, si les variants dans *TRRAP* induisent des défauts d'acétylation par TIP60, il serait intéressant de considérer l'acide valproïque comme piste de traitement.

Ainsi, l'utilisation de petites molécules inhibitrices générales pourrait permettre de rétablir des profils d'acétylation, de méthylation, ou de positionnement de nucléosomes spécifiques chez les individus porteurs de mutations dans *CHD3*, *TRRAP*, et *SMARCC2*, et ce particulièrement dans le cas de mutations gain de fonction type hypermorphes ou néomorphes. De plus, Chory et al. ont récemment montré qu'il était possible de développer des inhibiteurs de fonctions spécifiques de complexes de remodelage de la chromatine²⁶⁷. Ainsi, les auteurs ont pu

développer un inhibiteur spécifique de la répression génique dépendante du complexe BAF, qui n'affecte pas les capacités d'activation de la transcription des complexes BAF canoniques. Alors, le développement de ces molécules pourrait atténuer les symptômes causés par une sous-unité spécifique de chacun des complexes étudiés sans empêcher le reste du complexe de fonctionner correctement. Caractériser le mécanisme particulier par lequel le variant entraîne la maladie (perte ou gain de fonction) permettrait donc un traitement plus spécifique et avec potentiellement moins d'effets secondaires. Les molécules déjà identifiées pour le complexe BAF seraient des pistes thérapeutiques importantes pour tous les TDISS, dont celui causé par des variants dans *SMARCC2*. En identifiant le gène causateur de chaque syndrome, il devient donc possible de tester un traitement particulier au type d'enzyme qui est affecté (HAT, remodeleur de nucléosome). Ceci n'est possible qu'avec un diagnostic génétique car le diagnostic spécifique de chacun des syndromes sur le phénotype seulement n'est pas possible à cause de leur chevauchement.

Un traitement plus précis est aussi possible grâce à de nouvelles technologies de pointe dont la précision permet d'envisager leur utilisation *in vivo*. Des variants dans le gène *FMRI* sont causateurs du syndrome du X fragile, l'une des causes les plus fréquentes de DI. Il a été démontré que l'utilisation de la Cas9 enzymatiquement inactive (dCas9) fusionnée à la déméthylase Tet1 permet de cibler spécifiquement un complexe déméthylase à un gène cible, ici le gène *FMRI*, et de rétablir son expression²⁶⁰. En inhibant l'hyperméthylation des îlots CpG en amont du gène *FMRI*, son expression est réactivée de façon permanente dans les cellules, et est maintenue quand les cellules sont greffées dans des cerveaux murins. Dans le cas d'une perte de fonction hétérozygote, augmenter l'expression de l'allèle sauvage pourrait rétablir un fonctionnement normal de la protéine. Au-delà d'une thérapie génique visant à réparer les variants pathogéniques dans les gènes étudiés, il est donc aussi nécessaire de déterminer quelles sont les cibles précises des complexes associés à *TRRAP*, *CHD3*, et *SMARCC2*, et d'identifier quels sont les gènes responsables du phénotype observé. Avec ces informations, il serait possible de développer des outils utilisant la dCas9 fusionnée à des modificateurs de la chromatine pour rétablir des profils épigénétiques normaux dans le cerveau d'enfants atteints.

4.4 Perspectives et expériences futures

Nos études ont permis de montrer qu'il y avait une différence dans les profils d'expression génique entre les individus porteurs de variants dans les gènes *TRRAP*, *CHD3*, et *SMARCC2*. Dans les trois cas, nous pourrions confirmer nos résultats en utilisant des techniques d'inactivation génique (CRISPR-Cas9, Flox/Cre) ou de « knockdown » (siRNA/shRNA/ASO) pour supprimer chacune des protéines et confirmer la dérégulation des gènes ciblés dans les mêmes modèles cellulaires (Fibroblastes et LCLs) avec des contrôles isogéniques.

La fonction enzymatique de CHD3 étant connue, on a pu tester *in vitro* l'activité de la protéine sauvage ainsi que de certaines des mutations les plus récurrentes à l'aide de tests de l'activité ATPase, et de l'activité de remodelage de nucléosomes¹⁹⁷. Ceci a permis de démontrer que les variants testés dans *CHD3*, bien qu'ils causent un phénotype similaire, peuvent donner lieu à des activités enzymatiques variables, c'est-à-dire un gain ou une perte de fonction. Pourtant, un des variants testés ne montre pas de différence significative dans les deux essais. Alors, sachant que la totalité du complexe est nécessaire pour la fonction de CHD3, il serait intéressant de tester la capacité des protéines mutantes à s'assembler correctement dans le complexe NuRD par fixation du complexe suivi d'une analyse par spectrométrie de masse. Aussi, les variants dans *CHD3* pourraient affecter la localisation du complexe NuRD au niveau de la chromatine. Sachant que le complexe NuRD intègre prioritairement mais concurremment différentes hélicases CHD au cours du développement neural, et que chaque complexe NuRD (CHD3/4/5) est recruté à des niveaux distincts aux promoteurs de certains gènes selon le stade de différenciation des SCNs ou PNEs, il est possible que l'impact d'un variant ne soit détectable qu'à un certain stade de différenciation¹²⁵. Il est donc important de tester plusieurs stades de différenciation de neurones à la fois *in vitro* et *in vivo*. Sachant que le complexe NuRD est aussi impliqué dans la déacétylation¹³⁴, il serait intéressant d'évaluer si les variants dans *CHD3* affectent aussi l'activité des HDAC 1/2 par ChIPseq des marques d'acétylation d'histones. En effet, si l'activité HDAC est dépendante de la formation complète du complexe et que celle-ci est contrainte par une protéine CHD3 mutante, alors l'acétylation par le complexe NuRD pourrait être entravée.

Dans le cas des deux autres protéines, il n'y a pas de fonction enzymatique connue que l'on pourrait tester facilement *in vitro*. Dans le cas de SMARCC2, il serait possible d'évaluer sa présence dans le complexe BAF en exprimant les versions sauvage et mutantes puis en évaluant la composition du complexe en le fixant au formaldéhyde et en analysant son contenu par spectrométrie de masse, comme l'a été fait précédemment²²⁷. Avec cette technique, on pourrait caractériser la formation du complexe BAF au complet et déterminer s'il y a certaines sous-unités singulièrement affectées. La même étude pourrait être effectuée pour TRRAP afin d'évaluer ses interactions avec le complexe TIP60, mais aussi avec d'autres facteurs de transcription (cMYC, E2F). En effet, les variants dans *TRRAP* présentant un regroupement significatif dans des régions distinctes ce qui suggère que des fonctions différentes de la protéine sont affectées par chaque groupe de variants. Alors, il est probable que ces groupes affectent le site de liaison de partenaires d'interaction spécifiques ; expliquant les phénotypes différents observés chez ces individus.

Bibliographie

1. Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*. 1953;171(4356):737-738.
2. Pauling L, Corey RB. Structure of the nucleic acids. *Nature*. 1953;171(4347):346.
3. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860-921. doi:10.1038/35057062
4. Little PFR. Structure and function of the human genome. *Genome Res*. 2005;15(12):1759-1766. doi:10.1101/gr.4560905
5. Lyko F. The DNA methyltransferase family: A versatile toolkit for epigenetic regulation. *Nat Rev Genet*. 2018;19(2):81-92. doi:10.1038/nrg.2017.80
6. Wu SC, Zhang Y. Active DNA demethylation: Many roads lead to Rome. *Nature Reviews Molecular Cell Biology*. 2010;11(9):607-620. doi:10.1038/nrm2950
7. Legault L-M, Bertrand-Lehouillier V, McGraw S. Pre-implantation alcohol exposure and developmental programming of FASD: An epigenetic perspective. *Biochem Cell Biol*. 2018;96(2):117-130. doi:10.1139/bcb-2017-0141
8. Reik W. Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature*. 2007;447(7143):425-432. doi:10.1038/nature05918
9. Barlow DP. Genomic imprinting: A mammalian epigenetic discovery model. *Annu Rev Genet*. 2011;45:379-403. doi:10.1146/annurev-genet-110410-132459
10. Buck-Koehntop BA, Defossez P-A. On how mammalian transcription factors recognize methylated DNA. *Epigenetics*. 2013;8(2):131-137. doi:10.4161/epi.23632
11. Zou X, Ma W, Solov'yov IA, Chipot C, Schulten K. Recognition of methylated DNA through methyl-CpG binding domain proteins. *Nucleic Acids Res*. 2012;40(6):2747-2758. doi:10.1093/nar/gkr1057
12. Almouzni G, Cedar H. Maintenance of Epigenetic Information. *Cold Spring Harb Perspect Biol*. 2016;8(5):a019372. doi:10.1101/cshperspect.a019372
13. Prakash K, Fournier D. Evidence for the implication of the histone code in building the genome structure. *Biosystems*. 2018;164:49-59. doi:10.1016/j.biosystems.2017.11.005
14. Hirota A, Nakajima-Koyama M, Ashida Y, Nishida E. The nucleosome remodeling and deacetylase complex protein CHD4 regulates neural differentiation of mouse embryonic stem cells by down-regulating p53. *J Biol Chem*. 2019;294(1):195-209. doi:10.1074/jbc.RA118.004086

15. Talbert PB, Henikoff S. Histone variants on the move: Substrates for chromatin dynamics. *Nature Reviews Molecular Cell Biology*. 2017;18(2):115-126. doi:10.1038/nrm.2016.148
16. Lawrence M, Daujat S, Schneider R. Lateral Thinking: How Histone Modifications Regulate Gene Expression. *Trends Genet*. 2016;32(1):42-56. doi:10.1016/j.tig.2015.10.007
17. Yu Q, Zhang X, Bi X. Roles of chromatin remodeling factors in the formation and maintenance of heterochromatin structure. *J Biol Chem*. 2011;286(16):14659-14669. doi:10.1074/jbc.M110.183269
18. Saksouk N, Simboeck E, Déjardin J. Constitutive heterochromatin formation and transcription in mammals. *Epigenetics & Chromatin*. 2015;8(1):3. doi:10.1186/1756-8935-8-3
19. Mermoud JE, Rowbotham SP, Varga-Weisz PD. Keeping chromatin quiet: How nucleosome remodeling restores heterochromatin after replication. *Cell Cycle*. 2011;10(23):4017-4025. doi:10.4161/cc.10.23.18558
20. Kloc A, Martienssen R. RNAi, heterochromatin and the cell cycle. *Trends Genet*. 2008;24(10):511-517. doi:10.1016/j.tig.2008.08.002
21. Nakagawa C, Kawakita A, Fukada T, Sugimoto K. Live-cell imaging of HP1 α throughout the cell cycle of mouse C3H10T1/2 cells and rhythmical flickering of heterochromatin dots in interphase. *Biosci Biotechnol Biochem*. 2014;78(4):556-564. doi:10.1080/09168451.2014.893184
22. Sobecki M, Mrouj K, Camasses A, et al. The cell proliferation antigen Ki-67 organises heterochromatin. *Elife*. 2016;5:e13722. doi:10.7554/eLife.13722
23. Xing Y, Li WX. Heterochromatin components in germline stem cell maintenance. *Sci Rep*. 2015;5:17463. doi:10.1038/srep17463
24. Workman JL. Nucleosome displacement in transcription. *Genes Dev*. 2006;20(15):2009-2017. doi:10.1101/gad.1435706
25. Wigley DB, Bowman GD. A glimpse into chromatin remodeling. *Nature Structural & Molecular Biology*. 2017;24(6):498-500. doi:10.1038/nsmb.3415
26. Zhou CY, Johnson SL, Gamarra NI, Narlikar GJ. Mechanisms of ATP-Dependent Chromatin Remodeling Motors. *Annual Review of Biophysics*. 2016;45(1):153-181. doi:10.1146/annurev-biophys-051013-022819
27. Henikoff S, Smith MM. Histone Variants and Epigenetics. *Cold Spring Harbor Perspectives in Biology*. 2015;7(1). doi:10.1101/cshperspect.a019364

28. Weber CM, Henikoff S. Histone variants: Dynamic punctuation in transcription. *Genes Dev.* 2014;28(7):672-682. doi:10.1101/gad.238873.114
29. Lorch Y, Maier-Davis B, Kornberg RD. Chromatin remodeling by nucleosome disassembly in vitro. *PNAS.* 2006;103(9):3090-3093. doi:10.1073/pnas.0511050103
30. Volle C, Dalal Y. Histone variants: The tricksters of the chromatin world☆. *Curr Opin Genet Dev.* 2014;0:8-14. doi:10.1016/j.gde.2013.11.006
31. Bönisch C, Hake SB. Histone H2A variants in nucleosomes and chromatin: More or less stable? *Nucleic Acids Res.* 2012;40(21):10719-10741. doi:10.1093/nar/gks865
32. Turinetto V, Giachino C. Multiple facets of histone variant H2AX: A DNA double-strand-break marker with several biological functions. *Nucleic Acids Res.* 2015;43(5):2489-2498. doi:10.1093/nar/gkv061
33. Mariño-Ramírez L, Kann MG, Shoemaker BA, Landsman D. Histone structure and nucleosome stability. *Expert Rev Proteomics.* 2005;2(5):719-729. doi:10.1586/14789450.2.5.719
34. Starkova TY, Polyanichko AM, Artamonova TO, et al. Post-translational modifications of linker histone H1 variants in mammals. *Phys Biol.* 2017;14(1):016005. doi:10.1088/1478-3975/aa551a
35. Ausio J, Dong F, van Holde KE. Use of selectively trypsinized nucleosome core particles to analyze the role of the histone “tails” in the stabilization of the nucleosome. *J Mol Biol.* 1989;206(3):451-463.
36. Cosgrove MS, Boeke JD, Wolberger C. Regulated nucleosome mobility and the histone code. *Nat Struct Mol Biol.* 2004;11(11):1037-1043. doi:10.1038/nsmb851
37. Kahn TG, Dorafshan E, Schultheis D, et al. Interdependence of PRC1 and PRC2 for recruitment to Polycomb Response Elements. *Nucleic Acids Res.* 2016;44(21):10132-10149. doi:10.1093/nar/gkw701
38. Zhang K, Tang H, Huang L, et al. Identification of acetylation and methylation sites of histone H3 from chicken erythrocytes by high-accuracy matrix-assisted laser desorption ionization-time-of-flight, matrix-assisted laser desorption ionization-postsource decay, and nanoelectrospray ionization tandem mass spectrometry. *Anal Biochem.* 2002;306(2):259-269.
39. Sims Iii RJ, Reinberg D. Processing the H3K36me3 signature. *Nature Genetics.* 2009;41(3):270-271. doi:10.1038/ng0309-270
40. Chantalat S, Depaux A, Héry P, et al. Histone H3 trimethylation at lysine 36 is associated with constitutive and facultative heterochromatin. *Genome Res.* 2011;21(9):1426-1437. doi:10.1101/gr.118091.110

41. Li F, Wan M, Zhang B, et al. Bivalent Histone Modifications and Development. *Curr Stem Cell Res Ther*. 2018;13(2):83-90. doi:10.2174/1574888X12666170123144743
42. Minoux M, Holwerda S, Vitobello A, et al. Gene bivalency at Polycomb domains regulates cranial neural crest positional identity. *Science*. 2017;355(6332). doi:10.1126/science.aal2913
43. Farooq Z, Banday S, Pandita TK, Altaf M. The many faces of histone H3K79 methylation. *Mutat Res Rev Mutat Res*. 2016;768:46-52. doi:10.1016/j.mrrev.2016.03.005
44. Tropberger P, Pott S, Keller C, et al. Regulation of transcription through acetylation of H3K122 on the lateral surface of the histone octamer. *Cell*. 2013;152(4):859-872. doi:10.1016/j.cell.2013.01.032
45. Li M, Dong Q, Zhu B. Aurora Kinase B Phosphorylates Histone H3.3 at Serine 31 during Mitosis in Mammalian Cells. *J Mol Biol*. 2017;429(13):2042-2045. doi:10.1016/j.jmb.2017.01.016
46. Wilkins BJ, Rall NA, Ostwal Y, et al. A cascade of histone modifications induces chromatin condensation in mitosis. *Science*. 2014;343(6166):77-80. doi:10.1126/science.1244508
47. Götz M, Huttner WB. The cell biology of neurogenesis. *Nat Rev Mol Cell Biol*. 2005;6(10):777-788. doi:10.1038/nrml1739
48. Taverna E, Götz M, Huttner WB. The cell biology of neurogenesis: Toward an understanding of the development and evolution of the neocortex. *Annu Rev Cell Dev Biol*. 2014;30:465-502. doi:10.1146/annurev-cellbio-101011-155801
49. Uzquiano A, Gladwyn-Ng I, Nguyen L, et al. Cortical progenitor biology: Key features mediating proliferation versus differentiation. *J Neurochem*. 2018;146(5):500-525. doi:10.1111/jnc.14338
50. Bond AM, Ming G-L, Song H. Adult Mammalian Neural Stem Cells and Neurogenesis: Five Decades Later. *Cell Stem Cell*. 2015;17(4):385-395. doi:10.1016/j.stem.2015.09.003
51. Lim DA, Alvarez-Buylla A. Adult neural stem cells stake their ground. *Trends Neurosci*. 2014;37(10):563-571. doi:10.1016/j.tins.2014.08.006
52. Kriegstein A, Alvarez-Buylla A. The Glial Nature of Embryonic and Adult Neural Stem Cells. *Annu Rev Neurosci*. 2009;32:149-184. doi:10.1146/annurev.neuro.051508.135600
53. Florio M, Heide M, Pinson A, et al. Evolution and cell-type specificity of human-specific genes preferentially expressed in progenitors of fetal neocortex. Gleeson JG, ed. *ELife*. 2018;7:e32332. doi:10.7554/eLife.32332

54. Florio M, Huttner WB. Neural progenitors, neurogenesis and the evolution of the neocortex. *Development*. 2014;141(11):2182-2194. doi:10.1242/dev.090571
55. Kowalczyk T, Pontious A, Englund C, et al. Intermediate Neuronal Progenitors (Basal Progenitors) Produce Pyramidal–Projection Neurons for All Layers of Cerebral Cortex. *Cereb Cortex*. 2009;19(10):2439-2450. doi:10.1093/cercor/bhn260
56. Smart IHM, Dehay C, Giroud P, Berland M, Kennedy H. Unique morphological features of the proliferative zones and postmitotic compartments of the neural epithelium giving rise to striate and extrastriate cortex in the monkey. *Cereb Cortex*. 2002;12(1):37-53.
57. Hardwick LJA, Ali FR, Azzarelli R, Philpott A. Cell cycle regulation of proliferation versus differentiation in the central nervous system. *Cell Tissue Res*. 2015;359:187-200. doi:10.1007/s00441-014-1895-8
58. Ohnuma S, Harris WA. Neurogenesis and the Cell Cycle. *Neuron*. 2003;40(2):199-208. doi:10.1016/S0896-6273(03)00632-9
59. Calegari F, Haubensak W, Haffner C, Huttner WB. Selective lengthening of the cell cycle in the neurogenic subpopulation of neural progenitor cells during mouse brain development. *J Neurosci*. 2005;25(28):6533-6538. doi:10.1523/JNEUROSCI.0778-05.2005
60. Dobashi Y, Shoji M, Kitagawa M, Noguchi T, Kameya T. Simultaneous suppression of cdc2 and cdk2 activities induces neuronal differentiation of PC12 cells. *J Biol Chem*. 2000;275(17):12572-12580.
61. Wen S, Li H, Liu J. Dynamic signaling for neural stem cell fate determination. *Cell Adh Migr*. 2009;3(1):107-117.
62. Herrera A, Saade M, Menendez A, Marti E, Pons S. Sustained Wnt/ β -catenin signalling causes neuroepithelial aberrations through the accumulation of aPKC at the apical pole. *Nature Communications*. 2014;5:4168. doi:10.1038/ncomms5168
63. Woodhead GJ, Mutch CA, Olson EC, Chenn A. Cell-Autonomous β -Catenin Signaling Regulates Cortical Precursor Proliferation. *J Neurosci*. 2006;26(48):12620-12630. doi:10.1523/JNEUROSCI.3180-06.2006
64. Varela-Nallar L, Inestrosa NC. Wnt signaling in the regulation of adult hippocampal neurogenesis. *Front Cell Neurosci*. 2013;7. doi:10.3389/fncel.2013.00100
65. Munji RN, Choe Y, Li G, Siegenthaler JA, Pleasure SJ. Wnt Signaling Regulates Neuronal Differentiation of Cortical Intermediate Progenitors. *J Neurosci*. 2011;31(5):1676-1687. doi:10.1523/JNEUROSCI.5404-10.2011

66. Agathocleous M, Iordanova I, Willardsen MI, et al. A directional Wnt/ β -catenin-Sox2-proneural pathway regulates the transition from proliferation to differentiation in the *Xenopus* retina. *Development*. 2009;136(19):3289-3299. doi:10.1242/dev.040451
67. Belgacem YH, Hamilton AM, Shim S, Spencer KA, Borodinsky LN. The Many Hats of Sonic Hedgehog Signaling in Nervous System Development and Disease. *J Dev Biol*. 2016;4(4). doi:10.3390/jdb4040035
68. Rowitch DH, S-Jacques B, Lee SM, Flax JD, Snyder EY, McMahon AP. Sonic hedgehog regulates proliferation and inhibits differentiation of CNS precursor cells. *J Neurosci*. 1999;19(20):8954-8965.
69. Wechsler-Reya RJ, Scott MP. Control of Neuronal Precursor Proliferation in the Cerebellum by Sonic Hedgehog. *Neuron*. 1999;22(1):103-114. doi:10.1016/S0896-6273(00)80682-0
70. Contreras-Cornejo H, Saucedo-Correa G, Oviedo-Boyso J, et al. The CSL proteins, versatile transcription factors and context dependent corepressors of the notch signaling pathway. *Cell Div*. 2016;11. doi:10.1186/s13008-016-0025-2
71. Yoon K, Gaiano N. Notch signaling in the mammalian central nervous system: Insights from mouse mutants. *Nature Neuroscience*. 2005;8(6):709-715. doi:10.1038/nn1475
72. Taylor MK, Kelly Y, Morrison SJ. Physiological Notch signaling promotes gliogenesis in the developing peripheral and central nervous systems. *Development*. 2007;134(13):2435-2447. doi:10.1242/dev.005520
73. Kageyama R, Ohtsuka T, Kobayashi T. Roles of Hes genes in neural development. *Development, Growth & Differentiation*. 2008;50(s1):S97-S103. doi:10.1111/j.1440-169X.2008.00993.x
74. Lee H-Y, Kléber M, Hari L, et al. Instructive role of Wnt/beta-catenin in sensory fate specification in neural crest stem cells. *Science*. 2004;303(5660):1020-1023. doi:10.1126/science.1091611
75. Hirabayashi Y, Itoh Y, Tabata H, et al. The Wnt/ β -catenin pathway directs neuronal differentiation of cortical neural precursor cells. *Development*. 2004;131(12):2791-2801. doi:10.1242/dev.01165
76. Pöschl J, Grammel D, Dorostkar MM, Kretschmar HA, Schüller U. Constitutive activation of β -Catenin in neural progenitors results in disrupted proliferation and migration of neurons within the central nervous system. *Developmental Biology*. 2013;374(2):319-332. doi:10.1016/j.ydbio.2012.12.001
77. Bond AM, Bhalala OG, Kessler JA. The Dynamic Role of Bone Morphogenetic Proteins in Neural Stem Cell Fate and Maturation. *Dev Neurobiol*. 2012;72(7):1068-1084. doi:10.1002/dneu.22022

78. Panchision DM, Pickel JM, Studer L, et al. Sequential actions of BMP receptors control neural precursor cell production and fate. *Genes Dev.* 2001;15(16):2094-2110. doi:10.1101/gad.894701
79. Miller FD, Gauthier AS. Timing Is Everything: Making Neurons versus Glia in the Developing Cortex. *Neuron.* 2007;54(3):357-369. doi:10.1016/j.neuron.2007.04.019
80. Liu B, Ma A, Zhang F, et al. MAZ mediates the cross-talk between CT-1 and NOTCH1 signaling during gliogenesis. *Scientific Reports.* 2016;6:21534. doi:10.1038/srep21534
81. Barnabé-Heider F, Wasylnka JA, Fernandes KJL, et al. Evidence that Embryonic Neurons Regulate the Onset of Cortical Gliogenesis via Cardiotrophin-1. *Neuron.* 2005;48(2):253-265. doi:10.1016/j.neuron.2005.08.037
82. Fukuda S, Abematsu M, Mori H, et al. Potentiation of astrogliogenesis by STAT3-mediated activation of bone morphogenetic protein-Smad signaling in neural stem cells. *Mol Cell Biol.* 2007;27(13):4931-4937. doi:10.1128/MCB.02435-06
83. Zhou Z-D, Kumari U, Xiao Z-C, Tan E-K. Notch as a molecular switch in neural stem cells. *IUBMB Life.* 2010;62(8):618-623. doi:10.1002/iub.362
84. Bradl M, Lassmann H. Oligodendrocytes: Biology and pathology. *Acta Neuropathol.* 2010;119(1):37-53. doi:10.1007/s00401-009-0601-5
85. Davies JE, Miller RH. Local Sonic Hedgehog Signaling Regulates Oligodendrocyte Precursor Appearance in Multiple Ventricular Zone Domains in the Chick Metencephalon. *Developmental Biology.* 2001;233(2):513-525. doi:10.1006/dbio.2001.0224
86. Nery S, Wichterle H, Fishell G. Sonic hedgehog contributes to oligodendrocyte specification in the mammalian forebrain. *Development.* 2001;128(4):527-540.
87. Wang L-C, Almazan G. Role of Sonic Hedgehog Signaling in Oligodendrocyte Differentiation. *Neurochem Res.* 2016;41(12):3289-3299. doi:10.1007/s11064-016-2061-3
88. Zhou Q, Anderson DJ. The bHLH transcription factors OLIG2 and OLIG1 couple neuronal and glial subtype specification. *Cell.* 2002;109(1):61-73.
89. Meijer DH, Kane MF, Mehta S, et al. Separated at birth? The functional and molecular divergence of OLIG1 and OLIG2. *Nature Reviews Neuroscience.* 2012;13(12):819-831. doi:10.1038/nrn3386
90. Hu J-G, Lü H-Z, Wang Y-X, Bao M-S, Zhao B-M, Zhou J-S. BMP signaling mediates astrocyte differentiation of oligodendrocyte progenitor cells. *Tohoku J Exp Med.* 2010;222(3):195-200.

91. Mekki-Dauriac S, Agius E, Kan P, Cochard P. Bone morphogenetic proteins negatively control oligodendrocyte precursor specification in the chick spinal cord. *Development*. 2002;129(22):5117-5130.
92. Juliandi B, Abematsu M, Nakashima K. Chromatin remodeling in neural stem cell differentiation. *Curr Opin Neurobiol*. 2010;20(4):408-415. doi:10.1016/j.conb.2010.04.001
93. Sokpor G, Castro-Hernandez R, Rosenbusch J, Staiger JF, Tuoc T. ATP-Dependent Chromatin Remodeling During Cortical Neurogenesis. *Front Neurosci*. 2018;12. doi:10.3389/fnins.2018.00226
94. Hota SK, Bruneau BG. ATP-dependent chromatin remodeling during mammalian development. *Development*. 2016;143(16):2882-2897. doi:10.1242/dev.128892
95. Sokpor G, Xie Y, Rosenbusch J, Tuoc T. Chromatin Remodeling BAF (SWI/SNF) Complexes in Neural Development and Disorders. *Front Mol Neurosci*. 2017;10:243. doi:10.3389/fnmol.2017.00243
96. Bachmann C, Nguyen H, Rosenbusch J, et al. MSWI/SNF (BAF) Complexes Are Indispensable for the Neurogenesis and Development of Embryonic Olfactory Epithelium. *PLoS Genet*. 2016;12(9):e1006274. doi:10.1371/journal.pgen.1006274
97. Narayanan R, Pirouz M, Kerimoglu C, et al. Loss of BAF (mSWI/SNF) Complexes Causes Global Transcriptional and Chromatin State Changes in Forebrain Development. *Cell Rep*. 2015;13(9):1842-1854. doi:10.1016/j.celrep.2015.10.046
98. Tuoc T, Dere E, Radyushkin K, et al. Ablation of BAF170 in Developing and Postnatal Dentate Gyrus Affects Neural Stem Cell Proliferation, Differentiation, and Learning. *Mol Neurobiol*. 2017;54(6):4618-4635. doi:10.1007/s12035-016-9948-5
99. Ninkovic J, Steiner-Mezzadri A, Jawerka M, et al. The BAF complex interacts with Pax6 in adult neural progenitors to establish a neurogenic cross-regulatory transcriptional network. *Cell Stem Cell*. 2013;13(4):403-418. doi:10.1016/j.stem.2013.07.002
100. Georgala PA, Carr CB, Price DJ. The role of Pax6 in forebrain development. *Developmental Neurobiology*. 2011;71(8):690-709. doi:10.1002/dneu.20895
101. Lessard J, Wu JI, Ranish JA, et al. An essential switch in subunit composition of a chromatin remodeling complex during neural development. *Neuron*. 2007;55(2):201-215. doi:10.1016/j.neuron.2007.06.019
102. Staahl BT, Tang J, Wu W, et al. Kinetic analysis of npBAF to nBAF switching reveals exchange of SS18 with CREST and integration with neural developmental pathways. *J Neurosci*. 2013;33(25):10348-10361. doi:10.1523/JNEUROSCI.1258-13.2013

103. Vasileiou G, Ekici AB, Uebe S, et al. Chromatin-Remodeling-Factor ARID1B Represses Wnt/ β -Catenin Signaling. *American Journal of Human Genetics*. 2015;97(3):445. doi:10.1016/j.ajhg.2015.08.002
104. Middeljans E, Wan X, Jansen PW, Sharma V, Stunnenberg HG, Logie C. SS18 together with animal-specific factors defines human BAF-type SWI/SNF complexes. *PLoS ONE*. 2012;7(3):e33834. doi:10.1371/journal.pone.0033834
105. Nguyen H, Kerimoglu C, Pirouz M, et al. Epigenetic Regulation by BAF Complexes Limits Neural Stem Cell Proliferation by Suppressing Wnt Signaling in Late Embryonic Development. *Stem Cell Reports*. 2018;10(6):1734-1750. doi:10.1016/j.stemcr.2018.04.014
106. Zhan X, Shi X, Zhang Z, Chen Y, Wu JI. Dual role of Brg chromatin remodeling factor in Sonic hedgehog signaling during neural development. *Proc Natl Acad Sci U S A*. 2011;108(31):12758-12763. doi:10.1073/pnas.1018510108
107. Yoo AS, Staahl BT, Chen L, Crabtree GR. MicroRNA-mediated switching of chromatin-remodelling complexes in neural development. *Nature*. 2009;460(7255):642-646. doi:10.1038/nature08139
108. Yoo M, Choi K-Y, Kim J, et al. BAF53b, a Neuron-Specific Nucleosome Remodeling Factor, Is Induced after Learning and Facilitates Long-Term Memory Consolidation. *J Neurosci*. 2017;37(13):3686-3697. doi:10.1523/JNEUROSCI.3220-16.2017
109. Tuoc TC, Boretius S, Sansom SN, et al. Chromatin regulation by BAF170 controls cerebral cortical size and thickness. *Dev Cell*. 2013;25(3):256-269. doi:10.1016/j.devcel.2013.04.005
110. Tuoc TC, Narayanan R, Stoykova A. BAF chromatin remodeling complex: Cortical size regulation and beyond. *Cell Cycle*. 2013;12(18):2953-2959. doi:10.4161/cc.25999
111. Arlotta P, Molyneaux BJ, Jabaudon D, Yoshida Y, Macklis JD. Ctip2 Controls the Differentiation of Medium Spiny Neurons and the Establishment of the Cellular Architecture of the Striatum. *J Neurosci*. 2008;28(3):622-632. doi:10.1523/JNEUROSCI.2986-07.2008
112. Wiegreffe C, Simon R, Peschkes K, et al. Bcl11a (Ctip1) Controls Migration of Cortical Projection Neurons through Regulation of Sema3c. *Neuron*. 2015;87(2):311-325. doi:10.1016/j.neuron.2015.06.023
113. John A, Brylka H, Wiegreffe C, et al. Bcl11a is required for neuronal morphogenesis and sensory circuit formation in dorsal spinal cord development. *Development*. 2012;139(10):1831-1841. doi:10.1242/dev.072850

114. Choi K-Y, Yoo M, Han J-H. Toward understanding the role of the neuron-specific BAF chromatin remodeling complex in memory formation. *Experimental & Molecular Medicine*. 2015;47(4):e155. doi:10.1038/emm.2014.129
115. Aizawa H, Hu S-C, Bobb K, et al. Dendrite development regulated by CREST, a calcium-regulated transcriptional activator. *Science*. 2004;303(5655):197-202. doi:10.1126/science.1089845
116. Matsumoto Y, Otsuka F, Hino J, et al. Bone morphogenetic protein-3b (BMP-3b) inhibits osteoblast differentiation via Smad2/3 pathway by counteracting Smad1/5/8 signaling. *Mol Cell Endocrinol*. 2012;350(1):78-86. doi:10.1016/j.mce.2011.11.023
117. Yu Y, Chen Y, Kim B, et al. Olig2 Targets Chromatin Remodelers To Enhancers To Initiate Oligodendrocyte Differentiation. *Cell*. 2013;152(1-2):248-261. doi:10.1016/j.cell.2012.12.006
118. Bischof M, Weider M, Kuspert M, Nave K-A, Wegner M. Brg1-Dependent Chromatin Remodelling Is Not Essentially Required during Oligodendroglial Differentiation. *Journal of Neuroscience*. 2015;35(1):21-35. doi:10.1523/JNEUROSCI.1468-14.2015
119. Stanton BZ, Hodges C, Calarco JP, et al. Smarca4 ATPase mutations disrupt direct eviction of PRC1 from chromatin. *Nat Genet*. 2017;49(2):282-288. doi:10.1038/ng.3735
120. Kadoch C, Williams RT, Calarco JP, et al. Dynamics of BAF-Polycomb complex opposition on heterochromatin in normal and oncogenic states. *Nat Genet*. 2017;49(2):213-222. doi:10.1038/ng.3734
121. Basta J, Rauchman M. The Nucleosome Remodeling and Deacetylase (NuRD) Complex in Development and Disease. *Transl Res*. 2015;165(1):36-47. doi:10.1016/j.trsl.2014.05.003
122. Denslow SA, Wade PA. The human Mi-2/NuRD complex and gene regulation. *Oncogene*. 2007;26(37):5433-5438. doi:10.1038/sj.onc.1210611
123. López AJ, Wood MA. Role of nucleosome remodeling in neurodevelopmental and intellectual disability disorders. *Front Behav Neurosci*. 2015;9. doi:10.3389/fnbeh.2015.00100
124. Nitarska J, Smith JG, Sherlock WT, et al. A Functional Switch of NuRD Chromatin Remodeling Complex Subunits Regulates Mouse Cortical Development. *Cell Rep*. 2016;17(6):1683-1698. doi:10.1016/j.celrep.2016.10.022
125. Hoffmeister H, Fuchs A, Erdel F, et al. CHD3 and CHD4 form distinct NuRD complexes with different yet overlapping functionality. *Nucleic Acids Res*. 2017;45(18):10534-10554. doi:10.1093/nar/gkx711

126. Chudnovsky Y, Kim D, Zheng S, et al. ZFHX4 interacts with the NuRD core member CHD4 and regulates the glioblastoma tumor initiating cell state. *Cell Rep.* 2014;6(2):313-324. doi:10.1016/j.celrep.2013.12.032
127. Vestin A, Mills AA. The tumor suppressor Chd5 is induced during neuronal differentiation in the developing mouse brain. *Gene Expr Patterns.* 2013;13(8):482-489. doi:10.1016/j.gep.2013.09.003
128. Bishop B, Ho KK, Tyler K, et al. The chromatin remodeler chd5 is necessary for proper head development during embryogenesis of *Danio rerio*. *Biochim Biophys Acta.* 2015;1849(8):1040-1050. doi:10.1016/j.bbarm.2015.06.006
129. Montgomery RL, Hsieh J, Barbosa AC, Richardson JA, Olson EN. Histone deacetylases 1 and 2 control the progression of neural precursors to neurons during brain development. *PNAS.* 2009;106(19):7876-7881. doi:10.1073/pnas.0902750106
130. Tapias A, Wang Z-Q. Lysine Acetylation and Deacetylation in Brain Development and Neuropathies. *Genomics, Proteomics & Bioinformatics.* 2017;15(1):19-36. doi:10.1016/j.gpb.2016.09.002
131. Hagelkruys A, Lagger S, Krahmer J, et al. A single allele of Hdac2 but not Hdac1 is sufficient for normal mouse brain development in the absence of its paralog. *Development.* 2014;141(3):604-616. doi:10.1242/dev.100487
132. Kaji K, Caballero IM, MacLeod R, Nichols J, Wilson VA, Hendrich B. The NuRD component Mbd3 is required for pluripotency of embryonic stem cells. *Nat Cell Biol.* 2006;8(3):285-292. doi:10.1038/ncb1372
133. Baubec T, Ivánek R, Lienert F, Schübeler D. Methylation-dependent and -independent genomic targeting principles of the MBD protein family. *Cell.* 2013;153(2):480-492. doi:10.1016/j.cell.2013.03.011
134. Reynolds N, Salmon-Divon M, Dvinge H, et al. NuRD-mediated deacetylation of H3K27 facilitates recruitment of Polycomb Repressive Complex 2 to direct gene repression. *EMBO J.* 2012;31(3):593-605. doi:10.1038/emboj.2011.431
135. Mor N, Rais Y, Sheban D, et al. Neutralizing Gatad2a-Chd4-Mbd3/NuRD Complex Facilitates Deterministic Induction of Naive Pluripotency. *Cell Stem Cell.* 2018;23(3):412-425.e10. doi:10.1016/j.stem.2018.07.004
136. Murr R, Vaissière T, Sawan C, Shukla V, Herceg Z. Orchestration of chromatin-based processes: Mind the TRRAP. *Oncogene.* 2007;26(37):5358-5372. doi:10.1038/sj.onc.1210605
137. Herceg Z, Hulla W, Gell D, et al. Disruption of *Trrap* causes early embryonic lethality and defects in cell cycle progression. *Nature Genetics.* 2001;29(2):206-211. doi:10.1038/ng725

138. Knutson BA, Hahn S. Domains of Tra1 important for activator recruitment and transcription coactivator functions of SAGA and NuA4 complexes. *Mol Cell Biol.* 2011;31(4):818-831. doi:10.1128/MCB.00687-10
139. Flegel K, Grushko O, Bolin K, Griggs E, Buttitta L. Roles for the Histone Modifying and Exchange Complex NuA4 in Cell Cycle Progression in *Drosophila melanogaster*. *Genetics.* 2016;203(3):1265-1281. doi:10.1534/genetics.116.188581
140. Mo F, Zhuang X, Liu X, et al. Acetylation of Aurora B by TIP60 ensures accurate chromosomal segregation. *Nat Chem Biol.* 2016;12(4):226-232. doi:10.1038/nchembio.2017
141. Zhao L-J, Loewenstein PM, Green M. Identification of a panel of MYC and Tip60 co-regulated genes functioning primarily in cell cycle and DNA replication. *Genes Cancer.* 2018;9(3-4):101-113. doi:10.18632/genesandcancer.175
142. Shedden K, Cooper S. Analysis of cell-cycle gene expression in *Saccharomyces cerevisiae* using microarrays and multiple synchronization methods. *Nucleic Acids Res.* 2002;30(13):2920-2929.
143. Fischle W. Tip60-ing the balance in DSB repair. *Nature Cell Biology.* 2009;11(11):1279-1281. doi:10.1038/ncb1109-1279
144. Sun Y, Jiang X, Chen S, Fernandes N, Price BD. A role for the Tip60 histone acetyltransferase in the acetylation and activation of ATM. *PNAS.* 2005;102(37):13182-13187. doi:10.1073/pnas.0504211102
145. Tang Y, Luo J, Zhang W, Gu W. Tip60-dependent acetylation of p53 modulates the decision between cell-cycle arrest and apoptosis. *Mol Cell.* 2006;24(6):827-839. doi:10.1016/j.molcel.2006.11.021
146. Fazio TG, Huff JT, Panning B. An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. *Cell.* 2008;134(1):162-174. doi:10.1016/j.cell.2008.05.031
147. Giaimo BD, Ferrante F, Vallejo DM, et al. Histone variant H2A.Z deposition and acetylation directs the canonical Notch signaling response. *Nucleic Acids Res.* 2018;46(16):8197-8215. doi:10.1093/nar/gky551
148. Park JH, Sun X-J, Roeder RG. The SANT Domain of p400 ATPase Represses Acetyltransferase Activity and Coactivator Function of TIP60 in Basal p21 Gene Expression. *Molecular and Cellular Biology.* 2010;30(11):2750-2761. doi:10.1128/MCB.00804-09
149. Xu Y, Sun Y, Jiang X, et al. The p400 ATPase regulates nucleosome stability and chromatin ubiquitination during DNA repair. *The Journal of Cell Biology.* 2010;191(1):31-43. doi:10.1083/jcb.201001160

150. Chen PB, Hung J-H, Hickman TL, et al. Hdac6 regulates Tip60-p400 function in stem cells. *eLife*. 2013;2. doi:10.7554/eLife.01557
151. Acharya D, Hainer SJ, Yoon Y, et al. KAT-Independent Gene Regulation by Tip60 Promotes ESC Self-Renewal but Not Pluripotency. *Cell Reports*. 2017;19(4):671-679. doi:10.1016/j.celrep.2017.04.001
152. Rust K, Tiwari MD, Mishra VK, Grawe F, Wodarz A. Myc and the Tip60 chromatin remodeling complex control neuroblast maintenance and polarity in *Drosophila*. *EMBO J*. 2018;37(16). doi:10.15252/embj.201798659
153. Kim C-H, Kim J-W, Jang S-M, An J-H, Song K-H, Choi K-H. Transcriptional activity of paired homeobox Pax6 is enhanced by histone acetyltransferase Tip60 during mouse retina development. *Biochem Biophys Res Commun*. 2012;424(3):427-432. doi:10.1016/j.bbrc.2012.06.126
154. Tea JS, Luo L. The chromatin remodeling factor Bap55 functions through the TIP60 complex to regulate olfactory projection neuron dendrite targeting. *Neural Dev*. 2011;6:5. doi:10.1186/1749-8104-6-5
155. Xu S, Panikker P, Iqbal S, Elefant F. Tip60 HAT Action Mediates Environmental Enrichment Induced Cognitive Restoration. *PLoS One*. 2016;11(7). doi:10.1371/journal.pone.0159623
156. Panikker P, Xu S-J, Zhang H, et al. Restoring Tip60 HAT/HDAC2 Balance in the Neurodegenerative Brain Relieves Epigenetic Transcriptional Repression and Reinstates Cognition. *J Neurosci*. 2018;38(19):4569-4583. doi:10.1523/JNEUROSCI.2840-17.2018
157. Lorbeck M, Pirooznia K, Sarthi J, Zhu X, Elefant F. Microarray Analysis Uncovers a Role for Tip60 in Nervous System Function and General Metabolism. *PLoS One*. 2011;6(4). doi:10.1371/journal.pone.0018412
158. Bu P, Evrard YA, Lozano G, Dent SYR. Loss of Gcn5 Acetyltransferase Activity Leads to Neural Tube Closure Defects and Exencephaly in Mouse Embryos. *Mol Cell Biol*. 2007;27(9):3405-3416. doi:10.1128/MCB.00066-07
159. Lin W, Zhang Z, Srajer G, et al. Proper expression of the Gcn5 histone acetyltransferase is required for neural tube closure in mouse embryos. *Dev Dyn*. 2008;237(4):928-940. doi:10.1002/dvdy.21479
160. Martínez-Cerdeño V, Lemen JM, Chan V, et al. N-Myc and GCN5 Regulate Significantly Overlapping Transcriptional Programs in Neural Stem Cells. *PLoS One*. 2012;7(6). doi:10.1371/journal.pone.0039456
161. Kalkat M, Resetca D, Lourenco C, et al. MYC Protein Interactome Profiling Reveals Functionally Distinct Regions that Cooperate to Drive Tumorigenesis. *Molecular Cell*. 2018;72(5):836-848.e7. doi:10.1016/j.molcel.2018.09.031

162. Zhang N, Ichikawa W, Faiola F, Lo S-Y, Liu X, Martinez E. MYC interacts with the human STAGA coactivator complex via multivalent contacts with the GCN5 and TRRAP subunits. *Biochim Biophys Acta*. 2014;1839(5):395-405. doi:10.1016/j.bbagr.2014.03.017
163. Palhan VB, Chen S, Peng G-H, et al. Polyglutamine-expanded ataxin-7 inhibits STAGA histone acetyltransferase activity to produce retinal degeneration. *Proc Natl Acad Sci U S A*. 2005;102(24):8472-8477. doi:10.1073/pnas.0503505102
164. Xu W, Edmondson DG, Evrard YA, Wakamiya M, Behringer RR, Roth SY. Loss of Gcn5l2 leads to increased apoptosis and mesodermal defects during mouse development. *Nat Genet*. 2000;26(2):229-232. doi:10.1038/79973
165. Puttagunta R, Tedeschi A, S3ria MG, et al. PCAF-dependent epigenetic changes promote axonal regeneration in the central nervous system. *Nature Communications*. 2014;5:3527. doi:10.1038/ncomms4527
166. Lang SE, McMahon SB, Cole MD, Hearing P. E2F transcriptional activation requires TRRAP and GCN5 cofactors. *J Biol Chem*. 2001;276(35):32627-32634. doi:10.1074/jbc.M102067200
167. Memedula S, Belmont AS. Sequential Recruitment of HAT and SWI/SNF Components to Condensed Chromatin by VP16. *Current Biology*. 2003;13(3):241-246. doi:10.1016/S0960-9822(03)00048-4
168. Gause M, Eissenberg JC, MacRae AF, Dorsett M, Misulovin Z, Dorsett D. Nipped-A, the Tra1/TRRAP Subunit of the Drosophila SAGA and Tip60 Complexes, Has Multiple Roles in Notch Signaling during Wing Development. *Molecular and Cellular Biology*. 2006;26(6):2347-2359. doi:10.1128/MCB.26.6.2347-2359.2006
169. Mohan M, Herz H-M, Takahashi Y-H, et al. Linking H3K79 trimethylation to Wnt signaling through a novel Dot1-containing complex (DotCom). *Genes Dev*. 2010;24(6):574-589. doi:10.1101/gad.1898410
170. Finkbeiner MG, Sawan C, Ouzounova M, Murr R, Herceg Z. HAT cofactor TRRAP mediates beta-catenin ubiquitination on the chromatin and the regulation of the canonical Wnt pathway. *Cell Cycle*. 2008;7(24):3908-3914. doi:10.4161/cc.7.24.7354
171. Ceol CJ, Horvitz HR. A New Class of C. elegans synMuv Genes Implicates a Tip60/NuA4-like HAT Complex as a Negative Regulator of Ras Signaling. *Developmental Cell*. 2004;6(4):563-576. doi:10.1016/S1534-5807(04)00065-6
172. Sawan C, Hernandez-Vargas H, Murr R, et al. Histone Acetyltransferase Cofactor Trrap Maintains Self-Renewal and Restricts Differentiation of Embryonic Stem Cells. *STEM CELLS*. 2013;31(5):979-991. doi:10.1002/stem.1341

173. Tapias A, Zhou Z-W, Shi Y, et al. Trapp-Dependent Histone Acetylation Specifically Regulates Cell-Cycle Gene Transcription to Control Neural Progenitor Fate Decisions. *Cell Stem Cell*. 2014;14(5):632-643. doi:10.1016/j.stem.2014.04.001
174. Hardy S, Brand M, Mittler G, et al. TATA-binding Protein-free TAF-containing Complex (TFTC) and p300 Are Both Required for Efficient Transcriptional Activation. *J Biol Chem*. 2002;277(36):32875-32882. doi:10.1074/jbc.M205860200
175. Kato Y, Shi Y, He X. Neuralization of the *Xenopus* Embryo by Inhibition of p300/ CREB-Binding Protein Function. *J Neurosci*. 1999;19(21):9364-9373. doi:10.1523/JNEUROSCI.19-21-09364.1999
176. Yao T-P, Oh SP, Fuchs M, et al. Gene Dosage-Dependent Embryonic Development and Proliferation Defects in Mice Lacking the Transcriptional Integrator p300. *Cell*. 1998;93(3):361-372. doi:10.1016/S0092-8674(00)81165-4
177. Wang J, Weaver ICG, Gauthier-Fisher A, et al. CBP Histone Acetyltransferase Activity Regulates Embryonic Neural Differentiation in the Normal and Rubinstein-Taybi Syndrome Brain. *Developmental Cell*. 2010;18(1):114-125. doi:10.1016/j.devcel.2009.10.023
178. Nakashima K, Yanagisawa M, Arakawa H, et al. Synergistic Signaling in Fetal Brain by STAT3-Smad1 Complex Bridged by p300. *Science*. 1999;284(5413):479-482. doi:10.1126/science.284.5413.479
179. Zhang L, He X, Liu L, et al. Hdac3 Interaction with p300 Histone Acetyltransferase Regulates the Oligodendrocyte and Astrocyte Lineage Fate Switch. *Developmental Cell*. 2016;36(3):316-330. doi:10.1016/j.devcel.2016.01.002
180. Sun Y, Nadal-Vicens M, Misono S, et al. Neurogenin Promotes Neurogenesis and Inhibits Glial Differentiation by Independent Mechanisms. *Cell*. 2001;104(3):365-376. doi:10.1016/S0092-8674(01)00224-0
181. Mefford HC, Batshaw ML, Hoffman EP. Genomics, Intellectual Disability, and Autism. *The New England journal of medicine*. 2012;366(8):733. doi:10.1056/NEJMra1114194
182. Srivastava AK, Schwartz CE. Intellectual Disability and Autism Spectrum Disorders: Causal Genes and Molecular Mechanisms. *Neurosci Biobehav Rev*. 2014;46 Pt 2:161-174. doi:10.1016/j.neubiorev.2014.02.015
183. Moeschler JB. 3 - Neurodevelopmental Disabilities: Global Developmental Delay, Intellectual Disability, and Autism. In: Pyeritz RE, Korf BR, Grody WW, eds. *Emery and Rimoin's Principles and Practice of Medical Genetics and Genomics (Seventh Edition)*. Academic Press; 2019:61-79. doi:10.1016/B978-0-12-812536-6.00003-1
184. Baio J, Wiggins L, Christensen DL, et al. Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years—Autism and Developmental Disabilities Monitoring Network,

- 11 Sites, United States, 2014. *MMWR Surveill Summ.* 2018;67(6):1-23. doi:10.15585/mmwr.ss6706a1
185. Satterstrom FK, Kosmicki JA, Wang J, et al. Novel genes for autism implicate both excitatory and inhibitory cell lineages in risk. *BioRxiv.* December 2018:484113. doi:10.1101/484113
186. Lelieveld SH, Wiel L, Venselaar H, et al. Spatial Clustering of de Novo Missense Mutations Identifies Candidate Neurodevelopmental Disorder-Associated Genes. *The American Journal of Human Genetics.* 2017;101(3):478-484. doi:10.1016/j.ajhg.2017.08.004
187. Bögershausen N, Wollnik B. Mutational Landscapes and Phenotypic Spectrum of SWI/SNF-Related Intellectual Disability Disorders. *Front Mol Neurosci.* 2018;11. doi:10.3389/fnmol.2018.00252
188. Santen GWE, Aten E, Vulto-van Silfhout AT, et al. Coffin-Siris syndrome and the BAF complex: Genotype-phenotype study in 63 patients. *Hum Mutat.* 2013;34(11):1519-1528. doi:10.1002/humu.22394
189. Sousa SB, Hennekam RC, Nicolaides-Baraitser Syndrome International Consortium. Phenotype and genotype in Nicolaides-Baraitser syndrome. *Am J Med Genet C Semin Med Genet.* 2014;166C(3):302-314. doi:10.1002/ajmg.c.31409
190. Machol K, Rousseau J, Ehresmann S, et al. Expanding the Spectrum of BAF-Related Disorders: De Novo Variants in SMARCC2 Cause a Syndrome with Intellectual Disability and Developmental Delay. *The American Journal of Human Genetics.* 2019;104(1):164-178. doi:10.1016/j.ajhg.2018.11.007
191. Marfella CGA, Imbalzano AN. The Chd Family of Chromatin Remodelers. *Mutat Res.* 2007;618(1-2):30-40. doi:10.1016/j.mrfmmm.2006.07.012
192. Pilarowski GO, Vernon HJ, Applegate CD, et al. Missense variants in the chromatin remodeler CHD1 are associated with neurodevelopmental disability. *J Med Genet.* 2018;55(8):561-566. doi:10.1136/jmedgenet-2017-104759
193. Carvill GL, Heavin SB, Yendle SC, et al. Targeted resequencing in epileptic encephalopathies identifies de novo mutations in CHD2 and SYNGAP1. *Nat Genet.* 2013;45(7):825-830. doi:10.1038/ng.2646
194. Suls A, Jaehn JA, Kecskés A, et al. De novo loss-of-function mutations in CHD2 cause a fever-sensitive myoclonic epileptic encephalopathy sharing features with Dravet syndrome. *Am J Hum Genet.* 2013;93(5):967-975. doi:10.1016/j.ajhg.2013.09.017
195. Rauch A, Wieczorek D, Graf E, et al. Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: An exome sequencing study. *Lancet.* 2012;380(9854):1674-1682. doi:10.1016/S0140-6736(12)61480-9

196. Petersen AK, Streff H, Tokita M, Bostwick BL. The first reported case of an inherited pathogenic CHD2 variant in a clinically affected mother and daughter. *Am J Med Genet A*. 2018;176(7):1667-1669. doi:10.1002/ajmg.a.38835
197. Blok LS, Rousseau J, Twist J, et al. CHD3 helicase domain mutations cause a neurodevelopmental syndrome with macrocephaly and impaired speech and language. *Nature Communications*. 2018;9(1):4619. doi:10.1038/s41467-018-06014-6
198. Eising E, Carrion-Castillo A, Vino A, et al. A set of regulatory genes co-expressed in embryonic human brain is implicated in disrupted speech development. *Mol Psychiatry*. February 2018. doi:10.1038/s41380-018-0020-x
199. Sifrim A, Hitz M-P, Wilsdon A, et al. Distinct genetic architectures for syndromic and nonsyndromic congenital heart defects identified by exome sequencing. *Nat Genet*. 2016;48(9):1060-1065. doi:10.1038/ng.3627
200. Weiss K, Terhal PA, Cohen L, et al. De Novo Mutations in CHD4, an ATP-Dependent Chromatin Remodeler Gene, Cause an Intellectual Disability Syndrome with Distinctive Dysmorphisms. *Am J Hum Genet*. 2016;99(4):934-941. doi:10.1016/j.ajhg.2016.08.001
201. Pauli S, von Velsen N, Burfeind P, et al. CHD7 mutations causing CHARGE syndrome are predominantly of paternal origin. *Clin Genet*. 2012;81(3):234-239. doi:10.1111/j.1399-0004.2011.01701.x
202. Sanlaville D, Etchevers HC, Gonzales M, et al. Phenotypic spectrum of CHARGE syndrome in fetuses with CHD7 truncating mutations correlates with expression during human development. *J Med Genet*. 2006;43(3):211-217. doi:10.1136/jmg.2005.036160
203. Layman WS, McEwen DP, Beyer LA, et al. Defects in neural stem cell proliferation and olfaction in Chd7 deficient mice indicate a mechanism for hyposmia in human CHARGE syndrome. *Hum Mol Genet*. 2009;18(11):1909-1923. doi:10.1093/hmg/ddp112
204. Bernier R, Golzio C, Xiong B, et al. Disruptive CHD8 mutations define a subtype of autism early in development. *Cell*. 2014;158(2):263-276. doi:10.1016/j.cell.2014.06.017
205. O’Roak BJ, Vives L, Girirajan S, et al. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature*. 2012;485(7397):246-250. doi:10.1038/nature10989
206. Williams SR, Aldred MA, Der Kaloustian VM, et al. Haploinsufficiency of HDAC4 causes brachydactyly mental retardation syndrome, with brachydactyly type E, developmental delays, and behavioral problems. *Am J Hum Genet*. 2010;87(2):219-228. doi:10.1016/j.ajhg.2010.07.011
207. Chassaing N, Siani V, Carles D, et al. X-linked dominant chondrodysplasia with platyspondyly, distinctive brachydactyly, hydrocephaly, and microphthalmia. *Am J Med Genet A*. 2005;136A(4):307-312. doi:10.1002/ajmg.a.30570

208. Simon D, Laloo B, Barillot M, et al. A mutation in the 3'-UTR of the HDAC6 gene abolishing the post-transcriptional regulation mediated by hsa-miR-433 is linked to a new form of dominant X-linked chondrodysplasia. *Hum Mol Genet.* 2010;19(10):2015-2027. doi:10.1093/hmg/ddq083
209. Kaiser FJ, Ansari M, Braunholz D, et al. Loss-of-function HDAC8 mutations cause a phenotypic spectrum of Cornelia de Lange syndrome-like features, ocular hypertelorism, large fontanelle and X-linked inheritance. *Hum Mol Genet.* 2014;23(11):2888-2900. doi:10.1093/hmg/ddu002
210. Harakalova M, van den Boogaard M-J, Sinke R, et al. X-exome sequencing identifies a HDAC8 variant in a large pedigree with X-linked intellectual disability, truncal obesity, gynaecomastia, hypogonadism and unusual face. *J Med Genet.* 2012;49(8):539-543. doi:10.1136/jmedgenet-2012-100921
211. Lang B, Alrahbeni TMA, Clair DS, Blackwood DH, McCaig CD, Shen S. HDAC9 is implicated in schizophrenia and expressed specifically in post-mitotic neurons but not in adult neural stem cells. *Am J Stem Cells.* 2011;1(1):31-41.
212. Hobara T, Uchida S, Otsuki K, et al. Altered gene expression of histone deacetylases in mood disorder patients. *J Psychiatr Res.* 2010;44(5):263-270. doi:10.1016/j.jpsychires.2009.08.015
213. Pinto D, Delaby E, Merico D, et al. Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. *Am J Hum Genet.* 2014;94(5):677-694. doi:10.1016/j.ajhg.2014.03.018
214. Arboleda VA, Lee H, Dorrani N, et al. De novo nonsense mutations in KAT6A, a lysine acetyl-transferase gene, cause a syndrome including microcephaly and global developmental delay. *Am J Hum Genet.* 2015;96(3):498-506. doi:10.1016/j.ajhg.2015.01.017
215. Tham E, Lindstrand A, Santani A, et al. Dominant mutations in KAT6A cause intellectual disability with recognizable syndromic features. *Am J Hum Genet.* 2015;96(3):507-513. doi:10.1016/j.ajhg.2015.01.016
216. Campeau PM, Kim JC, Lu JT, et al. Mutations in KAT6B, encoding a histone acetyltransferase, cause Genitopatellar syndrome. *Am J Hum Genet.* 2012;90(2):282-289. doi:10.1016/j.ajhg.2011.11.023
217. Clayton-Smith J, O'Sullivan J, Daly S, et al. Whole-exome-sequencing identifies mutations in histone acetyltransferase gene KAT6B in individuals with the Say-Barber-Biesecker variant of Ohdo syndrome. *Am J Hum Genet.* 2011;89(5):675-681. doi:10.1016/j.ajhg.2011.10.008

218. Gannon T, Perveen R, Schlecht H, et al. Further delineation of the KAT6B molecular and phenotypic spectrum. *Eur J Hum Genet.* 2015;23(9):1165-1170. doi:10.1038/ejhg.2014.248
219. Simpson MA, Deshpande C, Dafou D, et al. De Novo Mutations of the Gene Encoding the Histone Acetyltransferase KAT6B Cause Genitopatellar Syndrome. *Am J Hum Genet.* 2012;90(2):290-294. doi:10.1016/j.ajhg.2011.11.024
220. Rubinstein JH, Taybi H. Broad thumbs and toes and facial abnormalities. A possible mental retardation syndrome. *Am J Dis Child.* 1963;105:588-608.
221. Hennekam RCM. Rubinstein–Taybi syndrome. *European Journal of Human Genetics.* 2006;14(9):981-985. doi:10.1038/sj.ejhg.5201594
222. Thienpont B, Béna F, Breckpot J, et al. Duplications of the critical Rubinstein-Taybi deletion region on chromosome 16p13.3 cause a novel recognisable syndrome. *J Med Genet.* 2010;47(3):155-161. doi:10.1136/jmg.2009.070573
223. Bartsch O, Labonté J, Albrecht B, et al. Two patients with EP300 mutations and facial dysmorphism different from the classic Rubinstein-Taybi syndrome. *Am J Med Genet A.* 2010;152A(1):181-184. doi:10.1002/ajmg.a.33153
224. Roelfsema JH, White SJ, Ariyürek Y, et al. Genetic heterogeneity in Rubinstein-Taybi syndrome: Mutations in both the CBP and EP300 genes cause disease. *Am J Hum Genet.* 2005;76(4):572-580. doi:10.1086/429130
225. Yan K, Rousseau J, Littlejohn RO, et al. Mutations in the Chromatin Regulator Gene BRPF1 Cause Syndromic Intellectual Disability and Deficient Histone Acetylation. *Am J Hum Genet.* 2017;100(1):91-104. doi:10.1016/j.ajhg.2016.11.011
226. Cogné B, Ehresmann S, Beauregard-Lacroix E, et al. Missense Variants in the Histone Acetyltransferase Complex Component Gene TRRAP Cause Autism and Syndromic Intellectual Disability. *Am J Hum Genet.* 2019;104(3):530-541. doi:10.1016/j.ajhg.2019.01.010
227. Mashtalir N, D’Avino AR, Michel BC, et al. Modular Organization and Assembly of SWI/SNF Family Chromatin Remodeling Complexes. *Cell.* 2018;175(5):1272-1288.e20. doi:10.1016/j.cell.2018.09.032
228. Monahan BJ, Villén J, Marguerat S, Bähler J, Gygi SP, Winston F. Fission yeast SWI/SNF and RSC complexes show compositional and functional differences from budding yeast. *Nat Struct Mol Biol.* 2008;15(8):873-880. doi:10.1038/nsmb.1452
229. Stern M, Jensen R, Herskowitz I. Five SWI genes are required for expression of the HO gene in yeast. *J Mol Biol.* 1984;178(4):853-868.

230. Boyer LA, Latek RR, Peterson CL. The SANT domain: A unique histone-tail-binding module? *Nature Reviews Molecular Cell Biology*. 2004;5(2):158-163. doi:10.1038/nrm1314
231. Boyer LA, Langer MR, Crowley KA, Tan S, Denu JM, Peterson CL. Essential role for the SANT domain in the functioning of multiple chromatin remodeling enzymes. *Mol Cell*. 2002;10(4):935-942.
232. Da G, Lenkart J, Zhao K, Shiekhata R, Cairns BR, Marmorstein R. Structure and function of the SWIRM domain, a conserved protein module found in chromatin regulatory complexes. *Proc Natl Acad Sci U S A*. 2006;103(7):2057-2062. doi:10.1073/pnas.0510949103
233. Aravind L, Iyer LM. The SWIRM domain: A conserved module found in chromosomal proteins points to novel chromatin-modifying activities. *Genome Biol*. 2002;3(8):research0039.1-research0039.7.
234. Ho L, Ronan JL, Wu J, et al. An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. *Proc Natl Acad Sci USA*. 2009;106(13):5181-5186. doi:10.1073/pnas.0812889106
235. Ragvin A, Valvatne H, Erdal S, et al. Nucleosome binding by the bromodomain and PHD finger of the transcriptional cofactor p300. *J Mol Biol*. 2004;337(4):773-788. doi:10.1016/j.jmb.2004.01.051
236. Weaver TM, Morrison EA, Musselman CA. Reading More than Histones: The Prevalence of Nucleic Acid Binding among Reader Domains. *Molecules*. 2018;23(10). doi:10.3390/molecules23102614
237. Bouazoune K, Mitterweger A, Längst G, et al. The dMi-2 chromodomains are DNA binding modules important for ATP-dependent nucleosome mobilization. *The EMBO Journal*. 2002;21(10):2430-2440. doi:10.1093/emboj/21.10.2430
238. McMahon SB, Van Buskirk HA, Dugan KA, Copeland TD, Cole MD. The novel ATM-related protein TRRAP is an essential cofactor for the c-Myc and E2F oncoproteins. *Cell*. 1998;94(3):363-374.
239. Jiang X, Sun Y, Chen S, Roy K, Price BD. The FATC Domains of PIKK Proteins Are Functionally Equivalent and Participate in the Tip60-dependent Activation of DNA-PKcs and ATM. *J Biol Chem*. 2006;281(23):15741-15746. doi:10.1074/jbc.M513172200
240. Bosotti R, Isacchi A, Sonnhammer EL. FAT: A novel domain in PIK-related kinases. *Trends Biochem Sci*. 2000;25(5):225-227.
241. Bourgey M, Dali R, Eveleigh R, et al. GenPipes: An open-source framework for distributed and scalable genomic analyses. *BioRxiv*. November 2018:459552. doi:10.1101/459552

242. Picard Tools By Broad Institute. <http://broadinstitute.github.io/picard/>. Accessed February 18, 2019.
243. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114-2120. doi:10.1093/bioinformatics/btu170
244. Dobin A, Davis CA, Schlesinger F, et al. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15-21. doi:10.1093/bioinformatics/bts635
245. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31(2):166-169. doi:10.1093/bioinformatics/btu638
246. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12). doi:10.1186/s13059-014-0550-8
247. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GOrilla: A tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics*. 2009;10(1):48. doi:10.1186/1471-2105-10-48
248. Mi H, Huang X, Muruganujan A, et al. PANTHER version 11: Expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res*. 2017;45(D1):D183-D189. doi:10.1093/nar/gkw1138
249. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nature Methods*. 2012;9(4):357-359. doi:10.1038/nmeth.1923
250. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078-2079. doi:10.1093/bioinformatics/btp352
251. Zhang Y, Liu T, Meyer CA, et al. Model-based Analysis of ChIP-Seq (MACS). *Genome Biology*. 2008;9(9):R137. doi:10.1186/gb-2008-9-9-r137
252. Heinz S, Benner C, Spann N, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell*. 2010;38(4):576-589. doi:10.1016/j.molcel.2010.05.004
253. Sharifi-Zarchi A, Gerovska D, Adachi K, et al. DNA methylation regulates discrimination of enhancers from promoters through a H3K4me1-H3K4me3 seesaw mechanism. *BMC Genomics*. 2017;18. doi:10.1186/s12864-017-4353-7
254. Aref-Eshghi E, Bend EG, Hood RL, et al. BAFopathies' DNA methylation epi-signatures demonstrate diagnostic utility and functional continuum of Coffin–Siris and Nicolaides–Baraitser syndromes. *Nature Communications*. 2018;9(1):4885. doi:10.1038/s41467-018-07193-y
255. Gabory A, Jammes H, Dandolo L. The H19 locus: Role of an imprinted non-coding RNA in growth and development. *Bioessays*. 2010;32(6):473-480. doi:10.1002/bies.200900170

256. Han C-L, Ge M, Liu Y-P, et al. LncRNA H19 contributes to hippocampal glial cell activation via JAK/STAT signaling in a rat model of temporal lobe epilepsy. *J Neuroinflammation*. 2018;15(1):103. doi:10.1186/s12974-018-1139-z
257. Kallen AN, Zhou X-B, Xu J, et al. The Imprinted H19 LncRNA Antagonizes Let-7 MicroRNAs. *Mol Cell*. 2013;52(1). doi:10.1016/j.molcel.2013.08.027
258. Koyama Y. Functional alterations of astrocytes in mental disorders: Pharmacological significance as a drug target. *Front Cell Neurosci*. 2015;9. doi:10.3389/fncel.2015.00261
259. Higashimori H, Morel L, Huth J, et al. Astroglial FMRP-dependent translational down-regulation of mGluR5 underlies glutamate transporter GLT1 dysregulation in the fragile X mouse. *Hum Mol Genet*. 2013;22(10):2041-2054. doi:10.1093/hmg/ddt055
260. Liu XS, Wu H, Krzisch M, et al. Rescue of Fragile X Syndrome Neurons by DNA Methylation Editing of the FMR1 Gene. *Cell*. 2018;172(5):979-992.e6. doi:10.1016/j.cell.2018.01.012
261. Yang Q, Feng B, Zhang K, et al. Excessive astrocyte-derived neurotrophin-3 contributes to the abnormal neuronal dendritic development in a mouse model of fragile X syndrome. *PLoS Genet*. 2012;8(12):e1003172. doi:10.1371/journal.pgen.1003172
262. Sauvageau M, Goff LA, Lodato S, et al. Multiple knockout mouse models reveal lincRNAs are required for life and brain development. *eLife*. 2013;2. doi:10.7554/eLife.01749
263. Li L, Zhuang Y, Zhao X, Li X. Long Non-coding RNA in Neuronal Development and Neurological Disorders. *Front Genet*. 2019;9. doi:10.3389/fgene.2018.00744
264. Guo W, Tsujimura K, Otsuka I M, et al. VPA alleviates neurological deficits and restores gene expression in a mouse model of Rett syndrome. *PLoS ONE*. 2014;9(6):e100215. doi:10.1371/journal.pone.0100215
265. Livide G, Patriarchi T, Amenduni M, et al. GluD1 is a common altered player in neuronal differentiation from both MECP2-mutated and CDKL5-mutated iPS cells. *Eur J Hum Genet*. 2015;23(2):195-201. doi:10.1038/ejhg.2014.81
266. Qin L, Ma K, Wang Z-J, et al. Social deficits in Shank3 -deficient mouse models of autism are rescued by histone deacetylase (HDAC) inhibition. *Nature Neuroscience*. 2018;21(4):564. doi:10.1038/s41593-018-0110-8
267. Chory EJ, Kirkland JG, Chang C-Y, et al. Inhibition of a Selective SWI/SNF Function Synergizes with ATR Inhibitors in Cancer Cell Killing. *BioRxiv*. June 2019. doi:10.1101/660456

Annexe I : Gènes différentiellement régulés dans les fibroblastes ou LCLs d'individus porteurs de variants dans *SMARCC2*, *TRRAP*, et *CHD3*.

| DEGs Fibros SMARCC2 | Symbol | log2FoldChange | padj | p.L610P | p.N134D | CTL - 1 | CTL - 2 | CTL - 3 | CTL - 4 |
|------------------------|----------|----------------|------------|---------|---------|---------|---------|---------|---------|
| ENSG00000002745 | WNT16 | -3.30036192 | 0.00159541 | 46 | 45 | 722 | 64 | 326 | 681 |
| ENSG00000002746 | HECW1 | -3.29197247 | 0.00016541 | 40 | 33 | 321 | 513 | 508 | 88 |
| ENSG00000003249 | DBNDD1 | -2.66798575 | 0.00632532 | 30 | 53 | 215 | 348 | 419 | 73 |
| ENSG00000021645 | NRXN3 | 4.09275461 | 0.00253113 | 14 | 259 | 2 | 6 | 17 | 7 |
| ENSG00000044524 | EPHA3 | 3.38170818 | 0.00215804 | 206 | 65 | 2 | 15 | 9 | 26 |
| ENSG00000049540 | ELN | 4.26166254 | 4.86E-09 | 98206 | 101102 | 1180 | 6950 | 7139 | 5512 |
| ENSG00000060140 | STYK1 | -2.22239111 | 0.00940238 | 16 | 41 | 139 | 83 | 122 | 188 |
| ENSG00000069122 | GPR116 | 4.62448736 | 0.00022242 | 37 | 37 | 1 | 3 | 1 | 1 |
| ENSG00000070669 | ASNS | 2.05314323 | 0.00059786 | 9155 | 4532 | 1406 | 1710 | 1457 | 2023 |
| ENSG00000077092 | RARB | 4.18878288 | 4.79E-07 | 379 | 314 | 10 | 10 | 9 | 47 |
| ENSG00000077943 | ITGA8 | 2.77747067 | 0.0007461 | 1024 | 3491 | 495 | 362 | 150 | 310 |
| ENSG00000078081 | LAMP3 | 3.92056377 | 0.00018085 | 48 | 58 | 2 | 2 | 2 | 8 |
| ENSG00000088756 | ARHGAP28 | 4.26589213 | 0.00028765 | 790 | 220 | 2 | 57 | 36 | 10 |
| ENSG00000099864 | PALM | 3.43653078 | 7.76E-05 | 2104 | 646 | 88 | 260 | 101 | 59 |
| ENSG00000101188 | NTSR1 | -6.86416353 | 0.00250752 | 3 | 4 | 308 | 75 | 1214 | 34 |
| ENSG00000101670 | LIPG | 3.50396185 | 0.00016289 | 118 | 211 | 29 | 18 | 6 | 5 |
| ENSG00000103449 | SALL1 | -2.09942114 | 0.00245475 | 87 | 125 | 237 | 552 | 515 | 513 |
| ENSG00000105664 | COMP | 3.5742729 | 0.00681206 | 19337 | 2312 | 883 | 676 | 1967 | 109 |
| ENSG00000106031 | HOXA13 | -4.3471595 | 0.00152098 | 2 | 15 | 83 | 323 | 271 | 15 |
| ENSG00000107623 | GDF10 | 4.78135512 | 0.00183208 | 161 | 4 | 2 | 3 | 6 | 1 |
| ENSG00000108602 | ALDH3A1 | -4.56332608 | 0.0022302 | 1 | 6 | 136 | 5 | 118 | 72 |
| ENSG00000110203 | FOLR3 | -3.91832415 | 0.00183208 | 12 | 55 | 1130 | 45 | 503 | 348 |
| ENSG00000111341 | MGP | 4.79344338 | 2.72E-05 | 522 | 1336 | 51 | 8 | 72 | 3 |
| ENSG00000111907 | TPD52L1 | 2.95989316 | 1.00E-07 | 1184 | 1512 | 198 | 108 | 246 | 141 |
| ENSG00000112559 | MDFI | 4.12337981 | 0.00115412 | 107 | 15 | 4 | 6 | 3 | 1 |
| ENSG00000114200 | BCHE | 3.45317124 | 0.00163013 | 49 | 66 | 4 | 12 | 4 | 1 |
| ENSG00000114270 | COL7A1 | -2.01919461 | 0.00317099 | 1461 | 1236 | 4445 | 9110 | 3813 | 4497 |

| | | | | | | | | | |
|------------------------|------------|-------------|------------|-------|-------|------|-------|-------|------|
| ENSG00000114315 | HES1 | 2.42146374 | 0.00104444 | 454 | 746 | 145 | 57 | 82 | 164 |
| ENSG00000115380 | EFEMP1 | 4.34777743 | 2.38E-09 | 88459 | 34036 | 1329 | 5257 | 1718 | 3728 |
| ENSG00000115602 | IL1RL1 | 3.11839394 | 0.00065623 | 99 | 66 | 10 | 4 | 6 | 18 |
| ENSG00000115604 | IL18R1 | 6.23405279 | 1.55E-15 | 271 | 143 | 3 | 1 | 5 | 2 |
| ENSG00000116183 | PAPPA2 | 3.37496485 | 0.00031247 | 1018 | 201 | 39 | 26 | 102 | 68 |
| ENSG00000117069 | ST6GALNAC5 | -3.05256591 | 0.00012899 | 113 | 143 | 735 | 666 | 2233 | 614 |
| ENSG00000120875 | DUSP4 | -3.50103623 | 0.00016289 | 90 | 157 | 1830 | 2738 | 523 | 502 |
| ENSG00000123892 | RAB38 | -3.39940305 | 0.0074207 | 11 | 18 | 13 | 269 | 185 | 145 |
| ENSG00000126218 | F10 | -2.16053474 | 0.00016289 | 282 | 386 | 2125 | 1191 | 1357 | 1300 |
| ENSG00000128655 | PDE11A | -3.33627995 | 0.00628205 | 10 | 15 | 155 | 13 | 171 | 166 |
| ENSG00000130487 | KLHDC7B | 4.06300748 | 0.00066953 | 97 | 20 | 2 | 2 | 3 | 7 |
| ENSG00000130595 | TNNT3 | 4.64385091 | 0.00327043 | 32 | 18 | 1 | 1 | 1 | 1 |
| ENSG00000130600 | H19 | 7.44905184 | 7.18E-43 | 3329 | 2350 | 11 | 14 | 30 | 10 |
| ENSG00000136928 | GABBR2 | -3.29594004 | 0.00754882 | 184 | 267 | 1808 | 311 | 5774 | 966 |
| ENSG00000137203 | TFAP2A | -3.63055504 | 0.00184572 | 25 | 84 | 655 | 1558 | 98 | 389 |
| ENSG00000137975 | CLCA2 | -3.69812152 | 0.00388446 | 5 | 19 | 330 | 18 | 116 | 159 |
| ENSG00000138207 | RBP4 | 3.29123016 | 0.00118098 | 65 | 28 | 5 | 5 | 4 | 5 |
| ENSG00000138378 | STAT4 | 2.18249009 | 0.00990217 | 791 | 564 | 168 | 51 | 186 | 192 |
| ENSG00000138678 | AGPAT9 | -2.70257015 | 0.00754882 | 12 | 40 | 67 | 118 | 235 | 257 |
| ENSG00000138792 | ENPEP | 2.55829424 | 0.00317099 | 108 | 160 | 11 | 39 | 27 | 14 |
| ENSG00000139269 | INHBE | 3.68588941 | 0.00483663 | 123 | 25 | 4 | 14 | 4 | 1 |
| ENSG00000139289 | PHLDA1 | -2.19097791 | 0.00103169 | 1942 | 2865 | 6736 | 15015 | 14208 | 7940 |
| ENSG00000139629 | GALNT6 | -2.37092358 | 0.00074044 | 873 | 392 | 1783 | 3103 | 4182 | 4019 |
| ENSG00000139970 | RTN1 | 3.74284077 | 0.00049122 | 30 | 298 | 11 | 20 | 9 | 9 |
| ENSG00000141576 | RNF157 | -3.18475778 | 0.0003665 | 14 | 29 | 91 | 223 | 360 | 108 |
| ENSG00000142910 | TINAGL1 | 3.6264122 | 0.00204375 | 634 | 3497 | 84 | 127 | 419 | 39 |
| ENSG00000144152 | FBLN7 | -2.24297753 | 0.0022302 | 204 | 211 | 461 | 1565 | 1019 | 884 |
| ENSG00000145936 | KCNMB1 | 5.10153241 | 0.00089237 | 6 | 97 | 1 | 2 | 1 | 2 |
| ENSG00000147509 | RGS20 | -2.27403186 | 0.0082856 | 45 | 102 | 153 | 493 | 429 | 347 |

| | | | | | | | | | |
|------------------------|-----------|-------------|------------|------|-------|------|------|------|------|
| ENSG00000148677 | ANKRD1 | 3.87191635 | 0.00010879 | 1581 | 11032 | 388 | 164 | 295 | 876 |
| ENSG00000149968 | MMP3 | -2.55929666 | 0.00127289 | 942 | 509 | 5274 | 1532 | 4128 | 6171 |
| ENSG00000152580 | IGSF10 | 3.33552198 | 2.99E-05 | 541 | 312 | 31 | 40 | 82 | 16 |
| ENSG00000154556 | SORBS2 | 3.503486 | 0.00088005 | 799 | 664 | 73 | 30 | 10 | 145 |
| ENSG00000156218 | ADAMTSL3 | 4.14683953 | 4.29E-05 | 66 | 58 | 1 | 6 | 6 | 1 |
| ENSG00000156466 | GDF6 | 4.0314539 | 0.0003906 | 2792 | 871 | 9 | 185 | 67 | 187 |
| ENSG00000156475 | PPP2R2B | -4.27696507 | 0.00068063 | 6 | 16 | 582 | 30 | 118 | 123 |
| ENSG00000162552 | WNT4 | 4.19703444 | 0.00115505 | 300 | 76 | 28 | 1 | 10 | 2 |
| ENSG00000162692 | VCAM1 | 3.98088999 | 9.35E-07 | 70 | 80 | 6 | 5 | 6 | 2 |
| ENSG00000162981 | FAM84A | -4.60287862 | 0.0003516 | 7 | 8 | 246 | 277 | 199 | 7 |
| ENSG00000164106 | SCRG1 | 5.43740346 | 9.30E-13 | 268 | 187 | 12 | 1 | 4 | 4 |
| ENSG00000164684 | ZNF704 | 4.05726655 | 2.80E-11 | 292 | 299 | 11 | 30 | 19 | 11 |
| ENSG00000164764 | SBSPON | 5.50461587 | 1.52E-05 | 15 | 212 | 4 | 4 | 1 | 1 |
| ENSG00000165061 | ZMAT4 | 4.9514538 | 2.03E-07 | 215 | 48 | 1 | 4 | 5 | 7 |
| ENSG00000165105 | RASEF | -3.37696765 | 2.93E-05 | 5 | 13 | 70 | 99 | 131 | 74 |
| ENSG00000165617 | DACT1 | 3.88532365 | 5.54E-12 | 6917 | 2644 | 305 | 249 | 295 | 445 |
| ENSG00000166473 | PKD1L2 | -3.46446472 | 0.00049122 | 5 | 8 | 38 | 100 | 108 | 41 |
| ENSG00000167941 | SOST | 5.87035674 | 0.00028536 | 6 | 111 | 1 | 1 | 1 | 1 |
| ENSG00000168209 | DDIT4 | 2.48556311 | 0.00118098 | 6789 | 1693 | 705 | 603 | 949 | 772 |
| ENSG00000170571 | EMB | -3.61579502 | 0.00787264 | 6 | 48 | 375 | 493 | 430 | 26 |
| ENSG00000170893 | TRH | 4.68182035 | 0.00023948 | 53 | 24 | 2 | 1 | 1 | 2 |
| ENSG00000172201 | ID4 | 3.25082245 | 0.00041126 | 2421 | 1320 | 50 | 108 | 310 | 318 |
| ENSG00000177133 | LINC00982 | 3.75488562 | 0.00089947 | 31 | 50 | 4 | 5 | 1 | 2 |
| ENSG00000177990 | DPY19L2 | -2.33847772 | 0.00835597 | 21 | 57 | 128 | 335 | 171 | 155 |
| ENSG00000178184 | PARD6G | -2.10132045 | 3.21E-05 | 153 | 170 | 614 | 833 | 587 | 738 |
| ENSG00000178662 | CSRNP3 | -3.02384485 | 0.00118098 | 6 | 9 | 70 | 54 | 87 | 33 |
| ENSG00000178919 | FOXE1 | 5.91886033 | 9.34E-11 | 89 | 637 | 3 | 12 | 6 | 3 |
| ENSG00000179772 | FOXS1 | 4.63226406 | 0.00068063 | 36 | 26 | 1 | 1 | 1 | 2 |
| ENSG00000180332 | KCTD4 | -3.85085201 | 0.00262301 | 3 | 18 | 62 | 29 | 202 | 313 |

| | | | | | | | | | |
|------------------------|------------------|-------------|------------|------|------|------|------|------|------|
| ENSG00000182389 | CACNB4 | 3.52222621 | 1.88E-07 | 270 | 270 | 42 | 11 | 18 | 23 |
| ENSG00000183770 | FOXL2 | -2.52842087 | 0.00618732 | 70 | 149 | 592 | 284 | 1233 | 418 |
| ENSG00000184005 | ST6GALNAC3 | 3.50733999 | 0.00552186 | 65 | 134 | 2 | 19 | 13 | 1 |
| ENSG00000185338 | SOCS1 | -2.56432084 | 0.00586859 | 37 | 51 | 167 | 442 | 336 | 96 |
| ENSG00000186352 | ANKRD37 | 2.35269762 | 0.00010891 | 287 | 566 | 67 | 93 | 93 | 81 |
| ENSG00000189056 | RELN | 3.71665191 | 0.00033884 | 846 | 541 | 16 | 141 | 19 | 35 |
| ENSG00000196436 | NPIP15 | -4.37068496 | 1.93E-07 | 3 | 5 | 75 | 87 | 74 | 95 |
| ENSG00000196562 | SULF2 | 2.3732263 | 0.00734394 | 6081 | 1820 | 816 | 334 | 990 | 910 |
| ENSG00000198597 | ZNF536 | -2.5986353 | 0.00521502 | 87 | 158 | 760 | 537 | 1414 | 257 |
| ENSG00000204176 | SYT15 | 3.536052 | 0.00032627 | 357 | 339 | 4 | 64 | 30 | 22 |
| ENSG00000213025 | COX20P1 | 4.75001715 | 0.00213289 | 5 | 143 | 3 | 6 | 1 | 1 |
| ENSG00000231924 | PSG1 | -3.77555308 | 0.00028765 | 22 | 55 | 188 | 140 | 709 | 1072 |
| ENSG00000236453 | AC003092.1 | -2.69487836 | 0.00740532 | 11 | 9 | 111 | 41 | 69 | 38 |
| ENSG00000254353 | RP11- 195E2.4 | -3.2684865 | 0.00157488 | 4 | 7 | 31 | 56 | 86 | 39 |
| ENSG00000257453 | RP11- 290L1.3 | -2.41740703 | 0.00964018 | 8 | 11 | 39 | 69 | 52 | 43 |
| ENSG00000261786 | RP4- 555D20.2 | -2.24265603 | 0.00658898 | 446 | 781 | 1798 | 2781 | 5290 | 1745 |
| ENSG00000268894 | PLCE1-AS1 | 3.86673204 | 0.00022678 | 73 | 299 | 7 | 6 | 7 | 31 |

| DEGs Fibros TRRAP | Symbol | log2FoldChange | padj | p.W186 6C | p.W186 6C | p.L805F | p.L805F | CTL-1 | CTL-1 | CTL-2 | CTL-2 |
|------------------------|---------------|----------------|----------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| ENSG00000152726 | FAM21B | 5.56390033 | 1.16E-45 | 205.060 967 | 199.404 112 | 203.646 753 | 161.220 346 | 2.12132 034 | 7.77817 459 | 2.82842 713 | 3.53553 391 |
| ENSG00000121361 | KCNJ8 | 3.95235454 | 1.43E-40 | 435.577 777 | 434.163 564 | 565.685 425 | 502.045 815 | 16.9705 628 | 43.1335 137 | 27.5771 645 | 37.4766 594 |
| ENSG00000162599 | NFIA | 2.28818792 | 3.81E-29 | 305.470 13 | 285.671 14 | 411.536 147 | 371.938 167 | 74.2462 12 | 67.8822 51 | 77.0746 392 | 62.2253 967 |
| ENSG00000115828 | QPCT | 2.24871311 | 1.39E-26 | 1067.73 124 | 1036.61 854 | 772.160 605 | 695.793 073 | 172.534 055 | 210.717 821 | 192.333 045 | 176.069 589 |
| ENSG00000150594 | ADRA2A | 2.70573871 | 1.52E-25 | 831.557 575 | 789.131 168 | 1279.86 327 | 1149.75 563 | 132.936 075 | 204.353 86 | 136.471 609 | 147.078 211 |
| ENSG00000137486 | ARRB1 | 2.26151662 | 9.78E-25 | 319.612 265 | 295.570 635 | 251.730 014 | 207.889 394 | 60.1040 764 | 56.5685 425 | 58.6898 628 | 48.7903 679 |
| ENSG00000131002 | TXLNG2P | -6.59120785 | 1.02E-22 | 1.41421 356 | 1.41421 356 | 1.41421 356 | 4.24264 069 | 176.776 695 | 236.880 772 | 180.312 229 | 224.152 85 |
| ENSG00000176435 | CLEC14A | 3.38285699 | 2.14E-21 | 183.847 763 | 173.948 268 | 203.646 753 | 190.918 831 | 10.6066 017 | 22.6274 17 | 10.6066 017 | 28.2842 713 |
| ENSG00000205664 | RP11-706O15.1 | 2.07800254 | 2.14E-20 | 110.308 658 | 110.308 658 | 110.308 658 | 125.865 007 | 24.0416 306 | 28.2842 713 | 27.5771 645 | 28.2842 713 |
| ENSG00000096696 | DSP | 2.72168931 | 4.61E-16 | 15562.0 06 | 18632.2 637 | 8297.19 097 | 8700.24 184 | 1584.62 63 | 2308.70 364 | 1781.20 198 | 2085.96 501 |
| ENSG00000255644 | RP11-59N23.1 | 3.59515798 | 4.62E-16 | 59.3969 696 | 91.9238 816 | 127.279 221 | 123.036 58 | 7.07106 781 | 9.89949 494 | 4.24264 069 | 12.0208 153 |
| ENSG00000120915 | EPHX2 | 2.61759349 | 6.83E-16 | 87.6812 409 | 100.409 163 | 131.521 861 | 123.036 58 | 24.0416 306 | 16.2634 56 | 20.5060 967 | 11.3137 085 |
| ENSG00000100077 | ADRBK2 | 2.45251219 | 1.77E-14 | 79.1959 595 | 79.1959 595 | 66.4680 374 | 69.2964 646 | 13.4350 288 | 9.89949 494 | 19.0918 831 | 11.3137 085 |
| ENSG00000169851 | PCDH7 | 3.08746279 | 5.41E-14 | 394.565 584 | 371.938 167 | 197.989 899 | 189.504 617 | 19.0918 831 | 44.5477 272 | 29.6984 848 | 42.4264 069 |
| ENSG00000069431 | ABCC9 | 4.22214792 | 6.10E-14 | 961.665 222 | 1023.89 062 | 1859.69 084 | 1723.92 633 | 26.1629 509 | 129.400 541 | 33.9411 255 | 108.894 444 |

| | | | | | | | | | | | |
|------------------------|---------------|-------------|----------|------------|------------|------------|------------|------------|------------|------------|------------|
| ENSG00000257022 | RP11-729I10.2 | 4.1351587 | 7.57E-13 | 41.0121933 | 28.2842713 | 57.9827561 | 46.6690476 | 1.41421356 | 3.53553391 | 1.41421356 | 3.53553391 |
| ENSG00000109743 | BST1 | 2.04036859 | 8.08E-13 | 899.439826 | 872.569768 | 1139.85613 | 1137.0277 | 291.327994 | 193.040151 | 330.925974 | 168.998521 |
| ENSG00000226259 | GTF2H2B | 3.51096145 | 2.20E-12 | 26.8700577 | 43.8406204 | 43.8406204 | 46.6690476 | 4.24264069 | 3.53553391 | 4.24264069 | 2.12132034 |
| ENSG00000144681 | STAC | 2.35888594 | 4.35E-12 | 360.624458 | 370.523953 | 608.111832 | 478.004184 | 110.308658 | 60.1040764 | 120.208153 | 63.6396103 |
| ENSG00000185008 | ROBO2 | -5.26223277 | 6.33E-12 | 7.07106781 | 11.3137085 | 15.5563492 | 12.7279221 | 113.844192 | 823.7794 | 115.258405 | 738.21948 |
| ENSG00000090539 | CHRD | 3.14345256 | 2.19E-11 | 101.823377 | 96.1665222 | 76.3675324 | 69.2964646 | 3.53553391 | 16.9705628 | 6.36396103 | 12.0208153 |
| ENSG00000261189 | RP3-512B11.3 | 2.71706595 | 3.83E-11 | 103.23759 | 103.23759 | 80.6101731 | 52.3259018 | 17.6776695 | 15.5563492 | 11.3137085 | 7.07106781 |
| ENSG00000132359 | RAP1GAP2 | 2.99999974 | 5.00E-11 | 36.7695526 | 32.5269119 | 39.5979798 | 38.1837662 | 7.07106781 | 4.94974747 | 3.53553391 | 2.82842713 |
| ENSG00000217236 | SP9 | 4.93073586 | 5.93E-11 | 533.158513 | 602.454978 | 96.1665222 | 62.2253967 | 7.07106781 | 13.4350288 | 12.0208153 | 9.89949494 |
| ENSG00000164161 | HHIP | 2.1324503 | 6.05E-11 | 48.0832611 | 66.4680374 | 52.3259018 | 74.9533188 | 15.5563492 | 13.4350288 | 13.4350288 | 12.7279221 |
| ENSG00000230623 | RP11-469A15.2 | 2.98935251 | 9.70E-11 | 55.1543289 | 38.1837662 | 66.4680374 | 31.1126984 | 6.36396103 | 4.94974747 | 7.77817459 | 4.94974747 |
| ENSG00000227825 | SLC9A7P1 | 2.10218042 | 2.02E-10 | 89.0954544 | 94.7523087 | 69.2964646 | 77.7817459 | 11.3137085 | 26.1629509 | 21.2132034 | 18.3847763 |
| ENSG00000234449 | RP11-706O15.3 | 4.53915701 | 3.19E-10 | 36.7695526 | 48.0832611 | 18.3847763 | 28.2842713 | 1.41421356 | 1.41421356 | 1.41421356 | 1.41421356 |
| ENSG00000152402 | GUCY1A2 | 2.23653665 | 3.28E-10 | 159.806133 | 111.722871 | 165.462987 | 142.83557 | 31.8198052 | 21.9203102 | 48.7903679 | 20.5060967 |
| ENSG00000196196 | HRCT1 | 2.09953569 | 5.77E-10 | 72.1248917 | 65.0538239 | 62.2253967 | 55.1543289 | 9.89949494 | 21.2132034 | 15.5563492 | 12.7279221 |
| ENSG00000175928 | LRRN1 | 5.36256718 | 8.95E-10 | 138.592929 | 213.546248 | 29.6984848 | 25.4558441 | 1.41421356 | 2.82842713 | 4.24264069 | 1.41421356 |

| | | | | | | | | | | | |
|-----------------------------|----------------|------------|----------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| ENSG000001 34247 | PTGFRN | 2.86391972 | 1.04E-09 | 1562.70 599 | 1455.22 576 | 3054.70 13 | 2534.27 07 | 391.030 05 | 178.190 909 | 464.569 155 | 148.492 424 |
| ENSG000001 27928 | GNGT1 | 2.5161635 | 1.26E-09 | 164.048 773 | 192.333 045 | 91.9238 816 | 101.823 377 | 20.5060 967 | 30.4055 916 | 13.4350 288 | 31.8198 052 |
| ENSG000001 98807 | PAX9 | 2.39854929 | 8.22E-09 | 39.5979 798 | 57.9827 561 | 35.3553 391 | 31.1126 984 | 9.19238 816 | 7.07106 781 | 7.77817 459 | 7.07106 781 |
| ENSG000001 24134 | KCNS1 | 2.19264511 | 8.73E-09 | 141.421 356 | 145.663 997 | 66.4680 374 | 76.3675 324 | 20.5060 967 | 26.8700 577 | 26.1629 509 | 20.5060 967 |
| ENSG000002 68241 | AC018470. 1 | 4.76249879 | 1.11E-08 | 287.085 353 | 308.298 557 | 53.7401 154 | 22.6274 17 | 7.07106 781 | 8.48528 137 | 3.53553 391 | 5.65685 425 |
| ENSG000001 51883 | PARP8 | 5.17592124 | 1.34E-08 | 466.690 476 | 449.719 913 | 57.9827 561 | 48.0832 611 | 12.7279 221 | 2.82842 713 | 10.6066 017 | 2.12132 034 |
| ENSG000002 21955 | SLC12A8 | 2.23731597 | 1.79E-08 | 295.570 635 | 321.026 479 | 229.102 597 | 207.889 394 | 72.1248 917 | 34.6482 323 | 84.1457 07 | 32.5269 119 |
| ENSG000001 06789 | CORO2A | 2.24051698 | 1.94E-08 | 113.137 085 | 142.835 57 | 65.0538 239 | 56.5685 425 | 19.7989 899 | 19.7989 899 | 19.0918 831 | 21.2132 034 |
| ENSG000001 13504 | SLC12A7 | 2.15560544 | 2.03E-08 | 379.009 235 | 408.707 72 | 182.433 55 | 164.048 773 | 62.9325 035 | 74.2462 12 | 59.3969 696 | 57.9827 561 |
| ENSG000002 40747 | KRBOX1 | 2.4027591 | 2.03E-08 | 49.4974 747 | 41.0121 933 | 48.0832 611 | 82.0243 866 | 13.4350 288 | 7.77817 459 | 13.4350 288 | 7.07106 781 |
| ENSG000001 05963 | ADAP1 | 2.2052581 | 2.29E-08 | 124.450 794 | 134.350 288 | 84.8528 137 | 50.9116 883 | 20.5060 967 | 26.1629 509 | 19.0918 831 | 19.7989 899 |
| ENSG000001 79841 | AKAP5 | 2.39231736 | 2.31E-08 | 59.3969 696 | 67.8822 51 | 65.0538 239 | 45.2548 34 | 14.8492 424 | 5.65685 425 | 16.9705 628 | 7.77817 459 |
| ENSG000001 82272 | B4GALNT4 | 2.83310258 | 2.57E-08 | 100.409 163 | 103.237 59 | 222.031 529 | 173.948 268 | 12.0208 153 | 25.4558 441 | 10.6066 017 | 36.0624 458 |
| ENSG000001 25378 | BMP4 | 2.21900975 | 4.40E-08 | 56.5685 425 | 48.0832 611 | 50.9116 883 | 45.2548 34 | 13.4350 288 | 6.36396 103 | 16.2634 56 | 7.07106 781 |
| ENSG000001 63701 | IL17RE | 2.27008924 | 4.44E-08 | 263.043 723 | 336.582 828 | 147.078 211 | 123.036 58 | 53.0330 086 | 38.8908 73 | 55.1543 289 | 33.2340 187 |
| ENSG000001 05825 | TFPI2 | 2.09406577 | 6.41E-08 | 4876.20 836 | 5594.62 885 | 2151.01 883 | 2322.13 867 | 717.713 383 | 1011.86 98 | 785.595 634 | 984.999 746 |

| | | | | | | | | | | | |
|-----------------------------|----------------|-------------|----------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| ENSG000001 23243 | ITIH5 | -4.13876045 | 1.13E-07 | 2.82842 713 | 1.41421 356 | 19.7989 899 | 12.7279 221 | 236.880 772 | 79.9030 663 | 256.679 762 | 74.2462 12 |
| ENSG000001 70421 | KRT8 | 2.55323495 | 1.37E-07 | 271.529 004 | 302.641 702 | 934.795 165 | 844.285 497 | 86.2670 273 | 123.036 58 | 80.6101 731 | 111.015 765 |
| ENSG000001 82013 | PNMAL1 | 2.58204501 | 1.80E-07 | 94.7523 087 | 100.409 163 | 250.315 801 | 253.144 228 | 18.3847 763 | 39.5979 798 | 21.2132 034 | 37.4766 594 |
| ENSG000002 35655 | H3F3AP4 | 2.18442467 | 2.36E-07 | 168.291 414 | 190.918 831 | 263.043 723 | 261.629 509 | 63.6396 103 | 24.7487 373 | 74.2462 12 | 31.8198 052 |
| ENSG000002 03721 | LINC00862 | 2.02783422 | 2.63E-07 | 62.2253 967 | 57.9827 561 | 52.3259 018 | 49.4974 747 | 16.9705 628 | 10.6066 017 | 19.7989 899 | 7.07106 781 |
| ENSG000001 71631 | P2RY6 | 3.45066047 | 2.80E-07 | 38.1837 662 | 43.8406 204 | 15.5563 492 | 18.3847 763 | 4.24264 069 | 2.12132 034 | 1.41421 356 | 2.82842 713 |
| ENSG000001 11341 | MGP | 3.18844446 | 2.88E-07 | 43.8406 204 | 57.9827 561 | 90.5096 68 | 84.8528 137 | 7.07106 781 | 4.24264 069 | 16.9705 628 | 2.12132 034 |
| ENSG000001 62620 | LRRIQ3 | 2.83007457 | 3.30E-07 | 18.3847 763 | 22.6274 17 | 25.4558 441 | 24.0416 306 | 1.41421 356 | 4.24264 069 | 3.53553 391 | 3.53553 391 |
| ENSG000000 67141 | NEO1 | 2.74127473 | 3.31E-07 | 7659.38 065 | 8582.86 211 | 2684.17 734 | 2572.45 447 | 488.610 786 | 1158.24 091 | 523.259 018 | 1045.10 382 |
| ENSG000001 18276 | B4GALT6 | 2.59484402 | 3.81E-07 | 295.570 635 | 295.570 635 | 115.965 512 | 121.622 366 | 45.9619 408 | 21.9203 102 | 49.4974 747 | 19.7989 899 |
| ENSG000002 37510 | AC008268. 2 | 3.06228367 | 4.18E-07 | 39.5979 798 | 22.6274 17 | 21.2132 034 | 16.9705 628 | 3.53553 391 | 2.12132 034 | 4.24264 069 | 2.12132 034 |
| ENSG000001 82993 | C12orf60 | 2.00000001 | 4.68E-07 | 49.4974 747 | 35.3553 391 | 76.3675 324 | 56.5685 425 | 16.9705 628 | 9.19238 816 | 14.8492 424 | 13.4350 288 |
| ENSG000002 68658 | LINC00664 | -2.0246616 | 4.80E-07 | 11.3137 085 | 8.48528 137 | 9.89949 494 | 11.3137 085 | 43.8406 204 | 47.3761 543 | 35.3553 391 | 40.3050 865 |
| ENSG000001 64530 | PI16 | 3.92439372 | 4.86E-07 | 111.722 871 | 104.651 804 | 1138.44 192 | 931.966 738 | 42.4264 069 | 33.2340 187 | 42.4264 069 | 32.5269 119 |
| ENSG000000 06432 | MAP3K9 | 3.56187788 | 5.50E-07 | 67.8822 51 | 69.2964 646 | 22.6274 17 | 15.5563 492 | 2.82842 713 | 5.65685 425 | 2.12132 034 | 4.24264 069 |
| ENSG000001 76485 | PLA2G16 | 2.53098493 | 1.20E-06 | 295.570 635 | 347.896 536 | 173.948 268 | 147.078 211 | 20.5060 967 | 58.6898 628 | 21.9203 102 | 65.7609 307 |

| | | | | | | | | | | | |
|------------------------|----------------|-------------|----------|------------|------------|------------|------------|------------|------------|------------|------------|
| ENSG00000116035 | VAX2 | 2.20328357 | 1.48E-06 | 76.3675324 | 80.6101731 | 43.8406204 | 46.6690476 | 13.4350288 | 17.6776695 | 16.9705628 | 5.65685425 |
| ENSG00000170571 | EMB | 8.38520896 | 1.90E-06 | 2026.56804 | 1981.3132 | 7.07106781 | 4.24264069 | 2.12132034 | 2.12132034 | 4.94974747 | 2.82842713 |
| ENSG00000186281 | GPAT2 | 3.08048974 | 2.22E-06 | 366.281313 | 432.74935 | 145.663997 | 149.906638 | 57.9827561 | 12.0208153 | 45.9619408 | 13.4350288 |
| ENSG00000173898 | SPTBN2 | 4.27462045 | 2.37E-06 | 183.847763 | 202.232539 | 25.4558441 | 12.7279221 | 3.53553391 | 3.53553391 | 7.07106781 | 7.77817459 |
| ENSG00000123689 | GOS2 | 3.82516477 | 2.54E-06 | 76.3675324 | 121.622366 | 14.1421356 | 18.3847763 | 3.53553391 | 4.24264069 | 2.82842713 | 5.65685425 |
| ENSG00000140459 | CYP11A1 | 3.19877912 | 2.57E-06 | 16.9705628 | 15.5563492 | 67.882251 | 42.4264069 | 3.53553391 | 4.24264069 | 4.24264069 | 3.53553391 |
| ENSG00000162645 | GBP2 | 2.26716605 | 2.65E-06 | 605.283405 | 540.229581 | 1077.63074 | 1139.85613 | 90.509668 | 272.943218 | 101.11627 | 234.052345 |
| ENSG00000272763 | RP11-357H14.17 | 5.02236336 | 3.26E-06 | 91.9238816 | 77.7817459 | 5.65685425 | 8.48528137 | 1.41421356 | 1.41421356 | 1.41421356 | 1.41421356 |
| ENSG00000164125 | FAM198B | 2.69603697 | 3.61E-06 | 3184.80894 | 3098.54192 | 704.278354 | 646.295598 | 274.357431 | 308.298557 | 291.327994 | 304.055916 |
| ENSG00000088881 | EBF4 | 2.19703675 | 3.63E-06 | 42.4264069 | 29.6984848 | 32.5269119 | 28.2842713 | 5.65685425 | 4.94974747 | 12.7279221 | 5.65685425 |
| ENSG00000230438 | RP11-420G6.4 | 3.23703824 | 4.07E-06 | 36.7695526 | 26.8700577 | 21.2132034 | 8.48528137 | 2.12132034 | 2.12132034 | 3.53553391 | 2.12132034 |
| ENSG00000135636 | DYSF | 4.32438385 | 4.24E-06 | 1111.57186 | 1255.82164 | 63.6396103 | 62.2253967 | 31.1126984 | 21.9203102 | 45.9619408 | 25.4558441 |
| ENSG00000111057 | KRT18 | 2.78783932 | 4.78E-06 | 510.531096 | 554.371717 | 1848.37713 | 1603.71818 | 79.1959595 | 265.87215 | 88.3883477 | 220.617316 |
| ENSG00000112294 | ALDH5A1 | 2.30281931 | 4.80E-06 | 183.847763 | 193.747258 | 79.1959595 | 73.5391052 | 17.6776695 | 37.4766594 | 16.9705628 | 35.3553391 |
| ENSG00000189058 | APOD | -2.69116085 | 5.11E-06 | 4.24264069 | 7.07106781 | 1.41421356 | 4.24264069 | 29.6984848 | 27.5771645 | 31.1126984 | 21.2132034 |
| ENSG00000196220 | SRGAP3 | 2.13565507 | 5.84E-06 | 66.4680374 | 42.4264069 | 70.7106781 | 41.0121933 | 12.0208153 | 9.19238816 | 21.9203102 | 7.07106781 |

| | | | | | | | | | | | |
|------------------------|---------------|-------------|----------|------------|------------|------------|------------|------------|------------|------------|------------|
| ENSG00000204963 | PCDHA7 | 2.79908682 | 8.19E-06 | 43.8406204 | 45.254834 | 15.5563492 | 18.3847763 | 2.82842713 | 7.07106781 | 3.53553391 | 4.24264069 |
| ENSG00000146411 | SLC2A12 | 2.10764078 | 9.13E-06 | 158.391919 | 152.735065 | 69.2964646 | 52.3259018 | 18.3847763 | 26.1629509 | 22.627417 | 33.2340187 |
| ENSG00000225410 | AC107016.1 | 2.44513944 | 9.26E-06 | 39.5979798 | 59.3969696 | 104.651804 | 115.965512 | 9.19238816 | 16.263456 | 7.77817459 | 25.4558441 |
| ENSG00000180044 | C3orf80 | 2.74322411 | 9.36E-06 | 12.7279221 | 22.627417 | 43.8406204 | 29.6984848 | 4.94974747 | 5.65685425 | 3.53553391 | 2.12132034 |
| ENSG00000168824 | NSG1 | 3.25633931 | 1.67E-05 | 312.541197 | 366.281313 | 84.8528137 | 87.6812409 | 38.890873 | 8.48528137 | 34.6482323 | 7.07106781 |
| ENSG00000224114 | RP11-343H5.4 | 5.34576883 | 2.13E-05 | 166.8772 | 169.705628 | 7.07106781 | 1.41421356 | 1.41421356 | 1.41421356 | 2.12132034 | 3.53553391 |
| ENSG00000158220 | ESYT3 | 2.30932797 | 2.17E-05 | 101.823377 | 76.3675324 | 31.1126984 | 32.5269119 | 8.48528137 | 16.9705628 | 10.6066017 | 12.7279221 |
| ENSG00000158825 | CDA | 2.58644611 | 2.74E-05 | 520.430591 | 620.839754 | 125.865007 | 108.894444 | 65.0538239 | 62.9325035 | 50.9116883 | 50.2045815 |
| ENSG00000248429 | RP11-597D13.9 | 2.38102546 | 2.89E-05 | 603.869191 | 579.827561 | 142.83557 | 135.764502 | 69.2964646 | 73.5391052 | 70.0035713 | 67.882251 |
| ENSG00000100504 | PYGL | 2.03053113 | 3.22E-05 | 4442.0448 | 4347.29249 | 1511.7943 | 1315.21861 | 552.957503 | 873.276875 | 591.848376 | 825.193614 |
| ENSG00000231584 | FAHD2CP | 2.24152658 | 3.41E-05 | 91.9238816 | 144.249783 | 70.7106781 | 50.9116883 | 27.5771645 | 8.48528137 | 27.5771645 | 12.0208153 |
| ENSG00000197653 | DNAH10 | 2.96729591 | 3.48E-05 | 159.806133 | 128.693434 | 1110.15765 | 1006.92006 | 74.246212 | 79.9030663 | 86.2670273 | 67.1751442 |
| ENSG00000166510 | CCDC68 | -3.90688597 | 3.53E-05 | 7.07106781 | 4.24264069 | 1.41421356 | 8.48528137 | 29.6984848 | 152.735065 | 11.3137085 | 124.450794 |
| ENSG00000180818 | HOXC10 | 7.35586129 | 3.58E-05 | 3412.49733 | 3859.38881 | 15.5563492 | 9.89949494 | 19.7989899 | 2.82842713 | 18.3847763 | 3.53553391 |
| ENSG00000132182 | NUP210 | 3.9205636 | 4.38E-05 | 135.764502 | 138.592929 | 14.1421356 | 11.3137085 | 8.48528137 | 2.82842713 | 2.12132034 | 6.36396103 |
| ENSG00000019186 | CYP24A1 | 7.00373657 | 5.82E-05 | 2.82842713 | 1.41421356 | 616.597113 | 468.104689 | 2.12132034 | 2.12132034 | 2.12132034 | 2.12132034 |

| | | | | | | | | | | | |
|-----------------------------|-------------------|------------|---------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| ENSG000001 17600 | LPPR4 | 2.73335435 | 5.88E-05 | 461.033 621 | 509.116 883 | 76.3675 324 | 82.0243 866 | 44.5477 272 | 43.1335 137 | 46.6690 476 | 35.3553 391 |
| ENSG000002 50159 | RP11- 381K20.2 | 2.30812199 | 6.25E-05 | 14.1421 356 | 12.7279 221 | 24.0416 306 | 22.6274 17 | 4.24264 069 | 3.53553 391 | 2.82842 713 | 4.24264 069 |
| ENSG000002 67650 | CTD- 2553C6.1 | 2.19347976 | 6.42E-05 | 60.8111 832 | 26.8700 577 | 142.835 57 | 118.793 939 | 21.2132 034 | 19.7989 899 | 16.9705 628 | 18.3847 763 |
| ENSG000001 97134 | ZNF257 | 3.64385421 | 6.46E-05 | 7.07106 781 | 8.48528 137 | 43.8406 204 | 46.6690 476 | 2.12132 034 | 1.41421 356 | 3.53553 391 | 1.41421 356 |
| ENSG000001 72575 | RASGRP1 | 2.75429995 | 6.48E-05 | 41.0121 933 | 28.2842 713 | 179.605 122 | 185.261 977 | 16.9705 628 | 20.5060 967 | 16.2634 56 | 10.6066 017 |
| ENSG000001 23388 | HOXC11 | 7.15985471 | 6.67E-05 | 654.780 879 | 756.604 256 | 2.82842 713 | 1.41421 356 | 2.12132 034 | 1.41421 356 | 4.24264 069 | 2.12132 034 |
| ENSG000002 06552 | KRBOX1- AS1 | 2.38827033 | 6.76E-05 | 16.9705 628 | 19.7989 899 | 31.1126 984 | 57.9827 561 | 7.77817 459 | 3.53553 391 | 6.36396 103 | 6.36396 103 |
| ENSG000001 10328 | GALNT18 | 2.23379707 | 7.37E-05 | 22.6274 17 | 29.6984 848 | 66.4680 374 | 60.8111 832 | 7.07106 781 | 15.5563 492 | 5.65685 425 | 9.89949 494 |
| ENSG000001 51623 | NR3C2 | 2.8160368 | 7.41E-05 | 154.149 278 | 145.663 997 | 26.8700 577 | 26.8700 577 | 12.0208 153 | 11.3137 085 | 17.6776 695 | 9.19238 816 |
| ENSG000001 45247 | OCIAD2 | 2.0108886 | 8.13E-05 | 578.413 347 | 586.898 628 | 179.605 122 | 159.806 133 | 93.3380 951 | 81.3172 798 | 103.944 697 | 94.7523 087 |
| ENSG000002 66066 | POLRMTP1 | 2.04064187 | 8.15E-05 | 18.3847 763 | 31.1126 984 | 24.0416 306 | 28.2842 713 | 5.65685 425 | 10.6066 017 | 3.53553 391 | 4.94974 747 |
| ENSG000002 28630 | HOTAIR | 7.15175996 | 8.50E-05 | 528.915 872 | 473.761 543 | 1.41421 356 | 1.41421 356 | 1.41421 356 | 1.41421 356 | 2.82842 713 | 1.41421 356 |
| ENSG000001 53064 | BANK1 | 2.72513976 | 8.66E-05 | 89.0954 544 | 98.9949 494 | 25.4558 441 | 15.5563 492 | 9.89949 494 | 8.48528 137 | 10.6066 017 | 5.65685 425 |
| ENSG000002 13225 | AC018804. 7 | 3.20789203 | 8.83E-05 | 145.663 997 | 145.663 997 | 18.3847 763 | 16.9705 628 | 5.65685 425 | 9.19238 816 | 9.89949 494 | 10.6066 017 |
| ENSG000000 21645 | NRXN3 | 4.1292799 | 9.51E-05 | 72.1248 917 | 63.6396 103 | 2.82842 713 | 9.89949 494 | 3.53553 391 | 2.12132 034 | 1.41421 356 | 1.41421 356 |
| ENSG000001 10693 | SOX6 | 3.11839386 | 0.000106 1 | 84.8528 137 | 110.308 658 | 15.5563 492 | 22.6274 17 | 12.7279 221 | 2.82842 713 | 7.07106 781 | 4.24264 069 |

| | | | | | | | | | | | |
|------------------------|---------------|------------|----------|---------|---------|---------|---------|---------|---------|---------|---------|
| ENSG00000080573 | COL5A3 | 2.20571293 | 0.000108 | 661.851 | 537.401 | 2709.63 | 2252.84 | 255.265 | 405.172 | 325.976 | 349.310 |
| | | | 12 | 947 | 154 | 319 | 221 | 548 | 186 | 226 | 75 |
| ENSG00000164106 | SCRG1 | 2.20645071 | 0.000110 | 12.7279 | 25.4558 | 46.6690 | 42.4264 | 7.77817 | 9.19238 | 4.94974 | 5.65685 |
| | | | 42 | 221 | 441 | 476 | 069 | 459 | 816 | 747 | 425 |
| ENSG00000260604 | RP1-140K8.5 | 2.15457701 | 0.000110 | 31.1126 | 48.0832 | 104.651 | 80.6101 | 19.0918 | 9.19238 | 22.6274 | 8.48528 |
| | | | 58 | 984 | 611 | 804 | 731 | 831 | 816 | 17 | 137 |
| ENSG00000184005 | ST6GALNA C3 | 3.18057124 | 0.000113 | 24.0416 | 11.3137 | 76.3675 | 80.6101 | 9.19238 | 3.53553 | 7.07106 | 1.41421 |
| | | | 16 | 306 | 085 | 324 | 731 | 816 | 391 | 781 | 356 |
| ENSG00000179300 | ZCCHC5 | 2.94110551 | 0.000113 | 11.3137 | 11.3137 | 56.5685 | 56.5685 | 5.65685 | 3.53553 | 3.53553 | 4.94974 |
| | | | 54 | 085 | 085 | 425 | 425 | 425 | 391 | 391 | 747 |
| ENSG00000228035 | RP4-663N10.1 | 2.13245009 | 0.000122 | 12.7279 | 28.2842 | 16.9705 | 22.6274 | 5.65685 | 5.65685 | 3.53553 | 3.53553 |
| | | | 58 | 221 | 713 | 628 | 17 | 425 | 425 | 391 | 391 |
| ENSG00000069011 | PITX1 | 3.43841295 | 0.000122 | 299.813 | 294.156 | 38.1837 | 42.4264 | 19.7989 | 3.53553 | 30.4055 | 8.48528 |
| | | | 96 | 275 | 421 | 662 | 069 | 899 | 391 | 916 | 137 |
| ENSG00000146197 | SCUBE3 | 2.54980651 | 0.000140 | 10581.1 | 10818.7 | 2037.88 | 2245.77 | 1428.35 | 743.876 | 1612.91 | 601.040 |
| | | | 6 | 459 | 338 | 174 | 114 | 57 | 334 | 057 | 764 |
| ENSG00000130203 | APOE | 2.70171449 | 0.000149 | 188.090 | 148.492 | 35.3553 | 28.2842 | 11.3137 | 24.0416 | 10.6066 | 15.5563 |
| | | | 52 | 404 | 424 | 391 | 713 | 085 | 306 | 017 | 492 |
| ENSG00000188064 | WNT7B | 5.82577679 | 0.000152 | 4.24264 | 7.07106 | 479.418 | 511.945 | 4.24264 | 2.82842 | 7.77817 | 2.82842 |
| | | | 28 | 069 | 781 | 398 | 31 | 069 | 713 | 459 | 713 |
| ENSG00000204642 | HLA-F | 2.01335853 | 0.000155 | 207.889 | 230.516 | 96.1665 | 79.1959 | 19.0918 | 57.2756 | 27.5771 | 48.0832 |
| | | | 04 | 394 | 811 | 222 | 595 | 831 | 493 | 645 | 611 |
| ENSG00000122367 | LDB3 | 3.01559591 | 0.000157 | 55.1543 | 53.7401 | 7.07106 | 15.5563 | 4.94974 | 2.12132 | 4.24264 | 4.94974 |
| | | | 23 | 289 | 154 | 781 | 492 | 747 | 034 | 069 | 747 |
| ENSG00000162551 | ALPL | 2.74075666 | 0.000158 | 21.2132 | 12.7279 | 79.1959 | 66.4680 | 7.77817 | 5.65685 | 9.89949 | 3.53553 |
| | | | 04 | 034 | 221 | 595 | 374 | 459 | 425 | 494 | 391 |
| ENSG00000254429 | CTD-2562J17.7 | 2.71881762 | 0.000161 | 43.8406 | 43.8406 | 9.89949 | 14.1421 | 4.24264 | 4.94974 | 2.12132 | 5.65685 |
| | | | 02 | 204 | 204 | 494 | 356 | 069 | 747 | 034 | 425 |
| ENSG00000250091 | DNAH100 S | 2.37808812 | 0.000163 | 63.6396 | 100.409 | 401.636 | 397.394 | 43.8406 | 54.4472 | 41.0121 | 45.9619 |
| | | | 81 | 103 | 163 | 652 | 011 | 204 | 222 | 933 | 408 |
| ENSG00000106483 | SFRP4 | 2.88752442 | 0.000164 | 36.7695 | 45.2548 | 15.5563 | 7.07106 | 2.82842 | 3.53553 | 5.65685 | 2.12132 |
| | | | 27 | 526 | 34 | 492 | 781 | 713 | 391 | 425 | 034 |

| | | | | | | | | | | | |
|------------------------|---------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| ENSG00000242611 | AC093627.8 | 2.7655338 | 0.00017014 | 12.7279221 | 14.1421356 | 12.7279221 | 8.48528137 | 2.12132034 | 1.41421356 | 1.41421356 | 2.12132034 |
| ENSG00000178882 | FAM101A | 2.82244892 | 0.00017482 | 1363.30187 | 1446.74047 | 212.132034 | 239.002092 | 160.513239 | 68.5893578 | 176.776695 | 55.1543289 |
| ENSG00000251151 | HOXC-AS3 | 6.29276922 | 0.00019307 | 357.796031 | 468.104689 | 1.41421356 | 4.24264069 | 4.94974747 | 1.41421356 | 2.12132034 | 2.12132034 |
| ENSG00000166145 | SPINT1 | 2.80735391 | 0.00019464 | 16.9705628 | 7.07106781 | 12.7279221 | 12.7279221 | 2.12132034 | 1.41421356 | 1.41421356 | 2.12132034 |
| ENSG00000272955 | KB-226F1.1 | 2.85414816 | 0.00019531 | 26.8700577 | 21.2132034 | 11.3137085 | 7.07106781 | 2.82842713 | 1.41421356 | 3.53553391 | 1.41421356 |
| ENSG00000158887 | MPZ | 2.30256258 | 0.00019572 | 59.3969696 | 62.2253967 | 18.3847763 | 16.9705628 | 5.65685425 | 7.77817459 | 8.48528137 | 9.89949494 |
| ENSG00000171224 | C10orf35 | 3.34169004 | 0.00019862 | 89.0954544 | 91.9238816 | 14.1421356 | 12.7279221 | 9.19238816 | 2.12132034 | 7.07106781 | 2.12132034 |
| ENSG00000166546 | BEAN1 | 2.46712549 | 0.00019905 | 12.7279221 | 8.48528137 | 18.3847763 | 26.8700577 | 3.53553391 | 2.82842713 | 2.82842713 | 2.82842713 |
| ENSG00000159263 | SIM2 | 3.01885823 | 0.00020053 | 94.7523087 | 94.7523087 | 18.3847763 | 9.89949494 | 4.94974747 | 9.19238816 | 4.94974747 | 7.77817459 |
| ENSG00000224149 | RP11-510C10.3 | 4.97727331 | 0.00020085 | 1.41421356 | 1.41421356 | 89.0954544 | 86.2670273 | 1.41421356 | 1.41421356 | 1.41421356 | 1.41421356 |
| ENSG00000118473 | SGIP1 | 2.57324556 | 0.00020269 | 1784.73752 | 1831.40656 | 265.87215 | 280.014285 | 201.525433 | 140.007143 | 197.989899 | 159.806133 |
| ENSG00000151572 | ANO4 | 2.36415626 | 0.00021778 | 87.6812409 | 118.793939 | 87.6812409 | 73.5391052 | 5.65685425 | 31.8198052 | 5.65685425 | 28.2842713 |
| ENSG00000196878 | LAMB3 | 2.8538112 | 0.00022982 | 1077.63074 | 999.848989 | 142.83557 | 100.409163 | 56.5685425 | 103.944697 | 67.882251 | 92.6309883 |
| ENSG00000166689 | PLEKHA7 | 2.05794729 | 0.0002433 | 56.5685425 | 63.6396103 | 22.627417 | 36.7695526 | 7.77817459 | 12.7279221 | 4.94974747 | 17.6776695 |
| ENSG00000234665 | RP11-262H14.3 | 2.69187688 | 0.00024436 | 21.2132034 | 19.7989899 | 7.07106781 | 11.3137085 | 3.53553391 | 1.41421356 | 2.12132034 | 2.12132034 |
| ENSG00000105409 | ATP1A3 | 3.24792663 | 0.00024792 | 127.279221 | 255.972655 | 25.4558441 | 21.2132034 | 6.36396103 | 19.0918831 | 4.94974747 | 14.8492424 |

| | | | | | | | | | | | |
|------------------------|------------|-------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| ENSG00000162624 | LHX8 | 6.46193723 | 0.0002556 | 1.41421356 | 1.41421356 | 427.092496 | 380.423448 | 2.12132034 | 2.12132034 | 3.53553391 | 1.41421356 |
| ENSG00000108932 | SLC16A6 | -4.87172341 | 0.00027387 | 4.24264069 | 4.24264069 | 1.41421356 | 2.82842713 | 4.24264069 | 183.847763 | 2.12132034 | 182.43355 |
| ENSG00000253958 | CLDN23 | 2.5670401 | 0.00027814 | 16.9705628 | 9.89949494 | 49.4974747 | 36.7695526 | 2.12132034 | 7.07106781 | 4.94974747 | 4.94974747 |
| ENSG00000108823 | SGCA | 2.05444759 | 0.00028623 | 21.2132034 | 9.89949494 | 22.627417 | 22.627417 | 4.94974747 | 4.94974747 | 5.65685425 | 2.82842713 |
| ENSG00000116833 | NR5A2 | 4.31563815 | 0.00032532 | 4.24264069 | 2.82842713 | 141.421356 | 175.362482 | 2.82842713 | 7.07106781 | 2.82842713 | 3.53553391 |
| ENSG00000095637 | SORBS1 | 3.12062878 | 0.00034918 | 9.89949494 | 14.1421356 | 113.137085 | 127.279221 | 9.19238816 | 8.48528137 | 6.36396103 | 6.36396103 |
| ENSG00000125851 | PCSK2 | 4.36256581 | 0.0003494 | 113.137085 | 84.8528137 | 4.24264069 | 1.41421356 | 3.53553391 | 1.41421356 | 2.82842713 | 2.12132034 |
| ENSG00000134532 | SOX5 | 2.8981192 | 0.00035326 | 16.9705628 | 25.4558441 | 9.89949494 | 5.65685425 | 2.82842713 | 1.41421356 | 2.12132034 | 1.41421356 |
| ENSG00000248458 | AL139147.1 | 2.33604903 | 0.00035441 | 98.9949494 | 77.7817459 | 16.9705628 | 24.0416306 | 10.6066017 | 12.7279221 | 10.6066017 | 9.19238816 |
| ENSG00000184845 | DRD1 | 3.42449606 | 0.00035459 | 7.07106781 | 11.3137085 | 53.7401154 | 72.1248917 | 7.07106781 | 1.41421356 | 3.53553391 | 1.41421356 |
| ENSG00000170689 | HOXB9 | 4.42998356 | 0.00035892 | 142.83557 | 125.865007 | 4.24264069 | 1.41421356 | 3.53553391 | 1.41421356 | 4.24264069 | 3.53553391 |
| ENSG00000163576 | EFHB | 2.19793901 | 0.00039836 | 11.3137085 | 18.3847763 | 12.7279221 | 12.7279221 | 3.53553391 | 2.82842713 | 2.82842713 | 2.82842713 |
| ENSG00000099998 | GGT5 | 2.89450613 | 0.00041065 | 11.3137085 | 12.7279221 | 89.0954544 | 91.9238816 | 7.07106781 | 5.65685425 | 5.65685425 | 9.19238816 |
| ENSG00000128285 | MCHR1 | 2.74322382 | 0.00041654 | 8.48528137 | 9.89949494 | 41.0121933 | 49.4974747 | 4.24264069 | 4.94974747 | 2.82842713 | 4.24264069 |
| ENSG00000077420 | APBB1IP | 2.60235882 | 0.00047851 | 1907.7741 | 1988.38427 | 783.474314 | 748.118975 | 50.9116883 | 405.172186 | 71.4177849 | 366.281313 |
| ENSG00000144488 | ESPNL | -2.02333116 | 0.00051949 | 39.5979798 | 35.3553391 | 12.7279221 | 9.89949494 | 117.379726 | 62.2253967 | 153.442172 | 63.6396103 |

| | | | | | | | | | | | |
|------------------------|---------------|------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| ENSG00000235180 | LINC00601 | 2.2016335 | 0.000546 44 | 24.0416 306 | 12.7279 221 | 18.3847 763 | 9.89949 494 | 5.65685 425 | 2.82842 713 | 2.82842 713 | 2.82842 713 |
| ENSG00000110002 | VWA5A | 2.0548841 | 0.000583 54 | 114.551 299 | 87.6812 409 | 271.529 004 | 246.073 16 | 64.3467 171 | 23.3345 238 | 66.4680 374 | 19.0918 831 |
| ENSG00000174611 | KY | 2.14201876 | 0.000617 55 | 14.1421 356 | 12.7279 221 | 31.1126 984 | 32.5269 119 | 8.48528 137 | 2.82842 713 | 4.24264 069 | 4.94974 747 |
| ENSG00000225864 | HCG4P11 | 2.67242485 | 0.000624 39 | 77.7817 459 | 93.3380 951 | 19.7989 899 | 25.4558 441 | 4.94974 747 | 11.3137 085 | 2.12132 034 | 15.5563 492 |
| ENSG00000169402 | RSPH10B2 | 2.5728892 | 0.000630 5 | 41.0121 933 | 49.4974 747 | 41.0121 933 | 36.7695 526 | 14.8492 424 | 1.41421 356 | 10.6066 017 | 1.41421 356 |
| ENSG00000163694 | RBM47 | 2.02769182 | 0.000652 38 | 100.409 163 | 111.722 871 | 74.9533 188 | 84.8528 137 | 6.36396 103 | 39.5979 798 | 11.3137 085 | 33.9411 255 |
| ENSG00000104435 | STMN2 | 2.86845332 | 0.000672 52 | 1146.92 72 | 1565.53 441 | 104.651 804 | 115.965 512 | 105.358 91 | 106.773 124 | 91.9238 816 | 97.5807 358 |
| ENSG00000198729 | PPP1R14C | 2.70043901 | 0.000674 77 | 50.9116 883 | 59.3969 696 | 9.89949 494 | 8.48528 137 | 4.94974 747 | 4.24264 069 | 4.94974 747 | 5.65685 425 |
| ENSG00000259886 | U82695.10 | 2.32192758 | 0.000699 92 | 14.1421 356 | 11.3137 085 | 18.3847 763 | 12.7279 221 | 2.12132 034 | 2.12132 034 | 5.65685 425 | 1.41421 356 |
| ENSG00000164197 | RNF180 | 2.96082838 | 0.000735 1 | 69.2964 646 | 62.2253 967 | 14.1421 356 | 8.48528 137 | 7.77817 459 | 1.41421 356 | 7.07106 781 | 3.53553 391 |
| ENSG00000186472 | PCLO | 3.64909046 | 0.000754 4 | 90.5096 68 | 94.7523 087 | 5.65685 425 | 4.24264 069 | 5.65685 425 | 1.41421 356 | 4.24264 069 | 4.24264 069 |
| ENSG00000124920 | MYRF | 2.01559675 | 0.000768 26 | 21.2132 034 | 14.1421 356 | 48.0832 611 | 48.0832 611 | 12.0208 153 | 5.65685 425 | 9.19238 816 | 5.65685 425 |
| ENSG00000141448 | GATA6 | 2.5823747 | 0.000796 46 | 53.7401 154 | 35.3553 391 | 364.867 099 | 333.754 401 | 40.3050 865 | 28.2842 713 | 38.1837 662 | 24.7487 373 |
| ENSG00000099251 | HSD17B7P 2 | 2.26303391 | 0.000801 66 | 16.9705 628 | 12.7279 221 | 12.7279 221 | 8.48528 137 | 2.82842 713 | 1.41421 356 | 3.53553 391 | 2.82842 713 |
| ENSG00000137857 | DUOX1 | 2.96593754 | 0.000811 08 | 203.646 753 | 173.948 268 | 19.7989 899 | 16.9705 628 | 12.7279 221 | 7.77817 459 | 22.6274 17 | 9.89949 494 |
| ENSG00000152078 | TMEM56 | 2.09686149 | 0.000867 9 | 67.8822 51 | 94.7523 087 | 35.3553 391 | 19.7989 899 | 16.2634 56 | 7.07106 781 | 19.0918 831 | 8.48528 137 |

| | | | | | | | | | | | |
|------------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| ENSG00000273340 | MICE | 2.32192755 | 0.00089807 | 12.7279221 | 24.0416306 | 9.89949494 | 9.89949494 | 2.82842713 | 3.53553391 | 1.41421356 | 3.53553391 |
| ENSG00000137699 | TRIM29 | 2.32930759 | 0.00090013 | 168.291414 | 181.019336 | 33.9411255 | 32.5269119 | 14.1421356 | 21.2132034 | 15.5563492 | 31.8198052 |
| ENSG00000170848 | PSG6 | 3.64769556 | 0.00090771 | 63.6396103 | 62.2253967 | 4.24264069 | 2.82842713 | 2.12132034 | 2.12132034 | 2.12132034 | 4.24264069 |
| ENSG00000150551 | LYPD1 | 2.40281593 | 0.00093264 | 38.1837662 | 46.6690476 | 229.102597 | 295.570635 | 23.3345238 | 40.3050865 | 19.7989899 | 31.8198052 |
| ENSG00000160190 | SLC37A1 | 2.18057179 | 0.00095668 | 14.1421356 | 9.89949494 | 11.3137085 | 12.7279221 | 2.12132034 | 2.82842713 | 2.12132034 | 3.53553391 |
| ENSG00000114115 | RBP1 | 2.23446479 | 0.00095807 | 9.89949494 | 14.1421356 | 22.627417 | 9.89949494 | 3.53553391 | 2.12132034 | 2.12132034 | 4.24264069 |
| ENSG00000101670 | LIPG | 3.1043354 | 0.00099633 | 7.07106781 | 15.5563492 | 86.2670273 | 73.5391052 | 7.07106781 | 2.12132034 | 9.89949494 | 2.12132034 |
| ENSG00000184292 | TACSTD2 | 5.28734231 | 0.00103673 | 2712.46161 | 2470.63109 | 33.9411255 | 29.6984848 | 50.9116883 | 7.07106781 | 67.1751442 | 9.19238816 |
| ENSG00000174343 | CHRNA9 | 2.54056771 | 0.00106843 | 33.9411255 | 36.7695526 | 12.7279221 | 7.07106781 | 5.65685425 | 2.12132034 | 5.65685425 | 2.12132034 |
| ENSG00000250451 | HOXC-AS1 | 2.98351065 | 0.0010725 | 53.7401154 | 52.3259018 | 9.89949494 | 7.07106781 | 6.36396103 | 2.12132034 | 5.65685425 | 1.41421356 |
| ENSG00000212710 | CTAGE1 | 3.74415775 | 0.00109536 | 42.4264069 | 46.6690476 | 4.24264069 | 1.41421356 | 1.41421356 | 1.41421356 | 2.12132034 | 2.12132034 |
| ENSG00000231360 | AL592284.1 | 2.8624951 | 0.00115168 | 21.2132034 | 24.0416306 | 4.24264069 | 7.07106781 | 2.82842713 | 1.41421356 | 2.12132034 | 1.41421356 |
| ENSG00000095587 | TLL2 | 2.47804713 | 0.00116108 | 137.178716 | 197.989899 | 31.1126984 | 19.7989899 | 10.6066017 | 19.0918831 | 20.5060967 | 19.0918831 |
| ENSG00000155629 | PIK3AP1 | 2.49476382 | 0.00117018 | 9.89949494 | 14.1421356 | 12.7279221 | 7.07106781 | 3.53553391 | 1.41421356 | 1.41421356 | 1.41421356 |
| ENSG00000105605 | CACNG7 | 3.63226609 | 0.00123499 | 11.3137085 | 9.89949494 | 176.776695 | 152.735065 | 15.5563492 | 2.12132034 | 9.19238816 | 1.41421356 |
| ENSG00000179388 | EGR3 | 3.26104505 | 0.00133263 | 12.7279221 | 18.3847763 | 408.70772 | 380.423448 | 19.7989899 | 19.0918831 | 23.3345238 | 23.3345238 |

| | | | | | | | | | | | |
|------------------------|-------------------|-------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| ENSG00000131459 | GFPT2 | 2.1805097 | 0.001374 36 | 295.570 635 | 291.327 994 | 2013.84 011 | 1704.12 734 | 243.244 733 | 232.638 131 | 260.922 402 | 212.839 141 |
| ENSG00000249790 | RP11- 20D14.6 | 3.42998603 | 0.001381 56 | 107.480 231 | 149.906 638 | 9.89949 494 | 7.07106 781 | 1.41421 356 | 12.7279 221 | 4.94974 747 | 6.36396 103 |
| ENSG00000223658 | AC011242. 6 | 2.03102662 | 0.001448 43 | 26.8700 577 | 16.9705 628 | 11.3137 085 | 11.3137 085 | 3.53553 391 | 2.12132 034 | 5.65685 425 | 4.94974 747 |
| ENSG00000135083 | CCNJL | 2.45169549 | 0.001475 5 | 28.2842 713 | 21.2132 034 | 33.9411 255 | 48.0832 611 | 8.48528 137 | 1.41421 356 | 12.7279 221 | 1.41421 356 |
| ENSG00000166831 | RBPMS2 | 4.07038512 | 0.001565 38 | 1.41421 356 | 1.41421 356 | 96.1665 222 | 79.1959 595 | 2.82842 713 | 2.82842 713 | 3.53553 391 | 1.41421 356 |
| ENSG00000099282 | TSPAN15 | 3.74542405 | 0.001571 21 | 77.7817 459 | 77.7817 459 | 2.82842 713 | 2.82842 713 | 5.65685 425 | 1.41421 356 | 2.12132 034 | 2.82842 713 |
| ENSG00000254963 | CTD- 2562J17.9 | 2.16227104 | 0.001608 89 | 28.2842 713 | 18.3847 763 | 9.89949 494 | 9.89949 494 | 2.82842 713 | 4.24264 069 | 2.12132 034 | 5.65685 425 |
| ENSG00000118526 | TCF21 | 2.38304744 | 0.001717 84 | 253.144 228 | 226.274 17 | 80.6101 731 | 70.7106 781 | 45.2548 34 | 9.89949 494 | 56.5685 425 | 9.19238 816 |
| ENSG00000085117 | CD82 | 2.09825312 | 0.001746 13 | 698.621 5 | 623.668 181 | 137.178 716 | 130.107 648 | 107.480 231 | 69.2964 646 | 139.300 036 | 55.1543 289 |
| ENSG00000171847 | FAM90A1 | 2.36923305 | 0.001831 53 | 15.5563 492 | 14.1421 356 | 7.07106 781 | 7.07106 781 | 1.41421 356 | 2.12132 034 | 2.82842 713 | 2.12132 034 |
| ENSG00000069188 | SDK2 | -3.86211225 | 0.001879 89 | 4.24264 069 | 2.82842 713 | 14.1421 356 | 12.7279 221 | 6.36396 103 | 285.671 14 | 4.94974 747 | 196.575 685 |
| ENSG00000133026 | MYH10 | 2.02231676 | 0.001889 38 | 2453.66 053 | 2470.63 109 | 14355.6 819 | 13705.1 436 | 1822.92 128 | 2288.90 465 | 2090.91 475 | 1916.96 648 |
| ENSG00000144063 | MALL | 2.24654646 | 0.001971 14 | 38.1837 662 | 59.3969 696 | 278.600 072 | 362.038 672 | 35.3553 391 | 42.4264 069 | 31.8198 052 | 45.9619 408 |
| ENSG00000259985 | RP11- 549B18.1 | 2.42884249 | 0.001973 78 | 18.3847 763 | 15.5563 492 | 9.89949 494 | 5.65685 425 | 4.24264 069 | 1.41421 356 | 1.41421 356 | 2.12132 034 |
| ENSG00000231205 | ZNF826P | 2.43812058 | 0.002067 61 | 12.7279 221 | 15.5563 492 | 45.2548 34 | 45.2548 34 | 2.82842 713 | 7.07106 781 | 1.41421 356 | 10.6066 017 |
| ENSG00000163884 | KLF15 | 2.28445306 | 0.002077 69 | 16.9705 628 | 14.1421 356 | 52.3259 018 | 50.9116 883 | 3.53553 391 | 12.7279 221 | 2.82842 713 | 8.48528 137 |

| | | | | | | | | | | | |
|------------------------|--------------------|-------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| ENSG00000112559 | MDFI | 3.89811622 | 0.0021128 | 1.41421356 | 1.41421356 | 52.3259018 | 60.8111832 | 2.12132034 | 1.41421356 | 2.82842713 | 1.41421356 |
| ENSG00000129467 | ADCY4 | 2.21065894 | 0.00211996 | 263.043723 | 287.085353 | 41.0121933 | 43.8406204 | 25.4558441 | 40.3050865 | 32.5269119 | 38.890873 |
| ENSG00000198963 | RORB | 3.9068863 | 0.00212246 | 50.9116883 | 52.3259018 | 1.41421356 | 1.41421356 | 2.82842713 | 1.41421356 | 1.41421356 | 1.41421356 |
| ENSG00000175591 | P2RY2 | 2.26303381 | 0.0021652 | 14.1421356 | 21.2132034 | 5.65685425 | 9.89949494 | 3.53553391 | 2.12132034 | 2.82842713 | 2.12132034 |
| ENSG00000149294 | NCAM1 | 2.335523 | 0.00218634 | 1019.64798 | 1117.22871 | 114.551299 | 161.220346 | 148.492424 | 95.4594155 | 151.320851 | 82.7314934 |
| ENSG00000128655 | PDE11A | 4.7498329 | 0.00233681 | 1057.83175 | 943.280446 | 8.48528137 | 7.07106781 | 14.8492424 | 25.4558441 | 14.1421356 | 20.5060967 |
| ENSG00000266283 | RP11-627G18.1 | 2.77137467 | 0.00238181 | 9.89949494 | 4.24264069 | 67.882251 | 57.9827561 | 6.36396103 | 4.94974747 | 3.53553391 | 5.65685425 |
| ENSG00000173535 | TNFRSF10C | 2.25434032 | 0.00239362 | 465.276262 | 485.075252 | 67.882251 | 57.9827561 | 56.5685425 | 58.6898628 | 62.9325035 | 47.3761543 |
| ENSG00000184058 | TBX1 | -3.06522249 | 0.00247806 | 15.5563492 | 8.48528137 | 28.2842713 | 18.3847763 | 12.7279221 | 299.813275 | 19.7989899 | 259.508189 |
| ENSG00000244588 | RAD21L1 | 2.58496147 | 0.00261439 | 24.0416306 | 18.3847763 | 14.1421356 | 2.82842713 | 1.41421356 | 4.24264069 | 1.41421356 | 2.82842713 |
| ENSG00000109265 | KIAA1211 | 2.97400311 | 0.00271525 | 38.1837662 | 31.1126984 | 2.82842713 | 5.65685425 | 2.82842713 | 2.12132034 | 3.53553391 | 1.41421356 |
| ENSG00000206341 | HLA-H | 2.02790609 | 0.0027467 | 147.078211 | 108.894444 | 49.4974747 | 63.6396103 | 36.0624458 | 7.07106781 | 37.4766594 | 9.89949494 |
| ENSG00000169085 | C8orf46 | 2.90689008 | 0.00279682 | 292.742207 | 271.529004 | 11.3137085 | 18.3847763 | 19.7989899 | 19.7989899 | 17.6776695 | 21.9203102 |
| ENSG00000053524 | MCF2L2 | 2.29698133 | 0.00290885 | 29.6984848 | 19.7989899 | 42.4264069 | 29.6984848 | 10.6066017 | 1.41421356 | 11.3137085 | 1.41421356 |
| ENSG00000198848 | CES1 | 4.08745842 | 0.00292395 | 1.41421356 | 1.41421356 | 130.107648 | 131.521861 | 5.65685425 | 1.41421356 | 7.07106781 | 1.41421356 |
| ENSG00000226210 | ABC7-42389800N19.1 | 2.31139737 | 0.00292753 | 79.1959595 | 67.882251 | 118.793939 | 120.208153 | 28.991378 | 3.53553391 | 40.3050865 | 4.94974747 |

| | | | | | | | | | | | |
|-----------------------------|-------------------|-------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| ENSG000001 15339 | GALNT3 | 3.26107177 | 0.002939 53 | 309.712 77 | 296.984 848 | 8.48528 137 | 8.48528 137 | 21.2132 034 | 12.0208 153 | 18.3847 763 | 13.4350 288 |
| ENSG000002 50133 | HOXC-AS2 | 3.44930503 | 0.003004 75 | 91.9238 816 | 98.9949 494 | 5.65685 425 | 4.24264 069 | 7.77817 459 | 1.41421 356 | 7.77817 459 | 1.41421 356 |
| ENSG000000 05471 | ABCB4 | 2.05614299 | 0.003067 66 | 96.1665 222 | 52.3259 018 | 24.0416 306 | 12.7279 221 | 12.0208 153 | 7.07106 781 | 14.1421 356 | 11.3137 085 |
| ENSG000002 42221 | PSG2 | 3.05595596 | 0.003138 08 | 509.116 883 | 475.175 757 | 28.2842 713 | 28.2842 713 | 12.0208 153 | 53.7401 154 | 12.7279 221 | 46.6690 476 |
| ENSG000000 55732 | MCOLN3 | 2.19264469 | 0.003143 77 | 29.6984 848 | 43.8406 204 | 7.07106 781 | 9.89949 494 | 4.94974 747 | 4.24264 069 | 5.65685 425 | 4.94974 747 |
| ENSG000001 89120 | SP6 | 2.0374745 | 0.003203 31 | 25.4558 441 | 22.6274 17 | 16.9705 628 | 45.2548 34 | 2.12132 034 | 6.36396 103 | 4.24264 069 | 14.1421 356 |
| ENSG000001 29465 | RIPK3 | 2.43706374 | 0.003485 84 | 205.060 967 | 243.244 733 | 25.4558 441 | 24.0416 306 | 15.5563 492 | 33.9411 255 | 16.9705 628 | 25.4558 441 |
| ENSG000001 14279 | FGF12 | 3.18762447 | 0.003595 46 | 28.2842 713 | 24.0416 306 | 1.41421 356 | 4.24264 069 | 1.41421 356 | 1.41421 356 | 1.41421 356 | 2.12132 034 |
| ENSG000001 32874 | SLC14A2 | 2.29545507 | 0.003676 99 | 8.48528 137 | 7.07106 781 | 15.5563 492 | 7.07106 781 | 1.41421 356 | 2.12132 034 | 2.12132 034 | 2.12132 034 |
| ENSG000001 42408 | CACNG8 | 3.60485851 | 0.003742 53 | 1.41421 356 | 1.41421 356 | 56.5685 425 | 43.8406 204 | 1.41421 356 | 2.82842 713 | 2.12132 034 | 2.12132 034 |
| ENSG000001 45423 | SFRP2 | 5.28076259 | 0.003872 85 | 4.24264 069 | 2.82842 713 | 1496.23 795 | 1135.61 349 | 21.2132 034 | 5.65685 425 | 33.9411 255 | 7.07106 781 |
| ENSG000001 97077 | KIAA1671 | 2.54056845 | 0.004030 68 | 376.180 808 | 400.222 438 | 53.7401 154 | 29.6984 848 | 53.7401 154 | 19.7989 899 | 55.8614 357 | 18.3847 763 |
| ENSG000001 35114 | OASL | -2.58495996 | 0.004037 24 | 4.24264 069 | 4.24264 069 | 2.82842 713 | 2.82842 713 | 16.9705 628 | 8.48528 137 | 48.7903 679 | 10.6066 017 |
| ENSG000002 54353 | RP11- 195E2.4 | 2.35049655 | 0.004211 27 | 5.65685 425 | 14.1421 356 | 22.6274 17 | 29.6984 848 | 4.94974 747 | 1.41421 356 | 6.36396 103 | 1.41421 356 |
| ENSG000001 57303 | SUSD3 | 2.9860593 | 0.004280 02 | 2.82842 713 | 7.07106 781 | 65.0538 239 | 70.7106 781 | 5.65685 425 | 3.53553 391 | 7.07106 781 | 2.12132 034 |
| ENSG000002 44649 | CTD- 2377D24.6 | 2.84799511 | 0.004352 25 | 26.8700 577 | 16.9705 628 | 4.24264 069 | 2.82842 713 | 2.12132 034 | 1.41421 356 | 2.12132 034 | 1.41421 356 |

| | | | | | | | | | | | |
|------------------------|---------------|-------------|-------------|------------|------------|------------|------------|------------|------------|------------|------------|
| ENSG00000174595 | KLF14 | 2.50589013 | 0.00458612 | 43.8406204 | 41.0121933 | 4.24264069 | 11.3137085 | 7.07106781 | 2.12132034 | 5.65685425 | 2.82842713 |
| ENSG00000247157 | RP11-434C1.1 | 2.6724241 | 0.00460731 | 32.5269119 | 28.2842713 | 4.24264069 | 7.07106781 | 1.41421356 | 5.65685425 | 1.41421356 | 2.82842713 |
| ENSG00000196503 | ARL9 | 2.31033955 | 0.0046444 | 11.3137085 | 4.24264069 | 36.7695526 | 35.3553391 | 4.94974747 | 3.53553391 | 6.36396103 | 2.82842713 |
| ENSG00000105989 | WNT2 | 2.34763389 | 0.00486221 | 132.936075 | 98.9949494 | 227.688384 | 181.019336 | 52.3259018 | 6.36396103 | 62.2253967 | 4.94974747 |
| ENSG00000186417 | GLDN | 2.49098572 | 0.00505169 | 12.7279221 | 4.24264069 | 72.1248917 | 57.9827561 | 4.94974747 | 4.94974747 | 9.19238816 | 7.07106781 |
| ENSG00000253125 | RP11-459E5.1 | 2.96962437 | 0.00509398 | 2.82842713 | 2.82842713 | 25.4558441 | 35.3553391 | 2.12132034 | 2.82842713 | 1.41421356 | 2.12132034 |
| ENSG00000144229 | THSD7B | 2.80735301 | 0.00515317 | 22.627417 | 8.48528137 | 5.65685425 | 2.82842713 | 1.41421356 | 1.41421356 | 1.41421356 | 1.41421356 |
| ENSG00000119508 | NR4A3 | 2.81950691 | 0.00521786 | 33.9411255 | 26.8700577 | 774.989032 | 681.650937 | 57.9827561 | 44.5477272 | 71.4177849 | 41.0121933 |
| ENSG00000261821 | CTD-2311M21.3 | 2.58496102 | 0.00525518 | 4.24264069 | 4.24264069 | 12.7279221 | 12.7279221 | 1.41421356 | 1.41421356 | 1.41421356 | 1.41421356 |
| ENSG00000107317 | PTGDS | 2.99999932 | 0.00530591 | 360.624458 | 265.87215 | 26.8700577 | 14.1421356 | 4.94974747 | 31.8198052 | 8.48528137 | 38.1837662 |
| ENSG00000146453 | PNLDC1 | 3.22238954 | 0.00539087 | 1.41421356 | 2.82842713 | 35.3553391 | 19.7989899 | 2.12132034 | 1.41421356 | 1.41421356 | 1.41421356 |
| ENSG00000185133 | INPP5J | 2.24392519 | 0.00558016 | 8.48528137 | 9.89949494 | 63.6396103 | 45.254834 | 5.65685425 | 4.24264069 | 9.89949494 | 7.07106781 |
| ENSG00000152503 | TRIM36 | 2.57511414 | 0.0056605 | 74.9533188 | 114.551299 | 8.48528137 | 8.48528137 | 7.77817459 | 11.3137085 | 10.6066017 | 4.94974747 |
| ENSG00000154146 | NRGN | 2.23162561 | 0.00567394 | 15.5563492 | 16.9705628 | 142.83557 | 120.208153 | 19.7989899 | 13.4350288 | 16.9705628 | 12.7279221 |
| ENSG00000101188 | NTSR1 | -5.02536897 | 0.005772713 | 2.82842713 | 1.41421356 | 8.48528137 | 18.3847763 | 2.12132034 | 505.581349 | 3.53553391 | 502.045815 |
| ENSG00000158560 | DYNC111 | 2.60108211 | 0.00585867 | 308.298557 | 407.293506 | 14.1421356 | 33.9411255 | 39.5979798 | 24.7487373 | 35.3553391 | 26.1629509 |

| | | | | | | | | | | | |
|------------------------|------------------|-------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| ENSG00000166796 | LDHC | 3.0324192 | 0.005922 16 | 4.24264 069 | 1.41421 356 | 21.2132 034 | 36.7695 526 | 2.82842 713 | 1.41421 356 | 2.12132 034 | 1.41421 356 |
| ENSG00000198133 | TMEM229 B | 2.30854881 | 0.005955 24 | 89.0954 544 | 93.3380 951 | 26.8700 577 | 18.3847 763 | 2.82842 713 | 21.9203 102 | 4.94974 747 | 16.2634 56 |
| ENSG00000170577 | SIX2 | 2.43812044 | 0.006038 04 | 7.07106 781 | 5.65685 425 | 49.4974 747 | 56.5685 425 | 7.07106 781 | 4.24264 069 | 4.24264 069 | 6.36396 103 |
| ENSG00000162817 | C1orf115 | 4.15889961 | 0.006085 76 | 469.518 903 | 434.163 564 | 9.89949 494 | 8.48528 137 | 19.0918 831 | 5.65685 425 | 24.0416 306 | 2.82842 713 |
| ENSG00000164251 | F2RL1 | 2.192942 | 0.006141 57 | 49.4974 747 | 49.4974 747 | 421.435 642 | 462.447 835 | 58.6898 628 | 48.0832 611 | 68.5893 578 | 39.5979 798 |
| ENSG00000119862 | LGALS1 | 2.36074732 | 0.006190 15 | 197.989 899 | 179.605 122 | 50.9116 883 | 50.9116 883 | 40.3050 865 | 6.36396 103 | 41.7193 001 | 4.94974 747 |
| ENSG00000230880 | RP11- 417J8.3 | 3.35754851 | 0.006279 57 | 1.41421 356 | 1.41421 356 | 32.5269 119 | 22.6274 17 | 1.41421 356 | 1.41421 356 | 1.41421 356 | 1.41421 356 |
| ENSG00000153898 | MCOLN2 | 2.29278108 | 0.006367 29 | 32.5269 119 | 21.2132 034 | 7.07106 781 | 8.48528 137 | 3.53553 391 | 1.41421 356 | 7.07106 781 | 2.12132 034 |
| ENSG00000105650 | PDE4C | 2.25375604 | 0.006553 67 | 32.5269 119 | 38.1837 662 | 12.7279 221 | 4.24264 069 | 2.82842 713 | 4.94974 747 | 2.82842 713 | 7.77817 459 |
| ENSG00000157423 | HYDIN | 2.16046408 | 0.006637 53 | 18.3847 763 | 21.2132 034 | 4.24264 069 | 9.89949 494 | 2.82842 713 | 2.82842 713 | 4.94974 747 | 1.41421 356 |
| ENSG00000196611 | MMP1 | -2.47399968 | 0.006665 88 | 1104.50 079 | 1336.43 182 | 8845.90 583 | 10011.2 178 | 8826.10 684 | 52730.3 669 | 10024.6 528 | 46747.5 364 |
| ENSG00000255282 | WTAPP1 | -2.46056953 | 0.007092 22 | 250.315 801 | 271.529 004 | 1880.90 404 | 2299.51 125 | 1986.26 295 | 11519.4 766 | 2262.74 17 | 10114.4 554 |
| ENSG00000099260 | PALMD | 2.94498545 | 0.007149 32 | 374.766 594 | 360.624 458 | 39.5979 798 | 25.4558 441 | 4.94974 747 | 44.5477 272 | 4.94974 747 | 49.4974 747 |
| ENSG00000197258 | EIF4BP6 | -3.004282 | 0.007303 98 | 7.07106 781 | 5.65685 425 | 7.07106 781 | 9.89949 494 | 4.24264 069 | 110.308 658 | 5.65685 425 | 118.086 833 |
| ENSG00000120949 | TNFRSF8 | 2.35363583 | 0.007318 72 | 4.24264 069 | 5.65685 425 | 11.3137 085 | 11.3137 085 | 1.41421 356 | 2.12132 034 | 1.41421 356 | 1.41421 356 |
| ENSG00000121005 | CRISPLD1 | 2.79354747 | 0.007529 69 | 36.7695 526 | 29.6984 848 | 5.65685 425 | 1.41421 356 | 2.82842 713 | 1.41421 356 | 3.53553 391 | 2.82842 713 |

| | | | | | | | | | | | |
|------------------------|---------------|-------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| ENSG00000180053 | NKX2-6 | -2.87183873 | 0.00777831 | 4.24264069 | 4.24264069 | 12.7279221 | 14.1421356 | 126.572114 | 8.48528137 | 116.672619 | 7.07106781 |
| ENSG00000247416 | RP11-629G13.1 | 2.48542585 | 0.0080128 | 38.1837662 | 31.1126984 | 4.24264069 | 5.65685425 | 4.24264069 | 2.12132034 | 5.65685425 | 2.12132034 |
| ENSG00000183196 | CHST6 | 3.32192513 | 0.00811104 | 1.41421356 | 1.41421356 | 62.2253967 | 48.0832611 | 4.24264069 | 2.12132034 | 3.53553391 | 1.41421356 |
| ENSG00000227487 | NCAM1-AS1 | 2.07972695 | 0.00837301 | 35.3553391 | 73.5391052 | 15.5563492 | 7.07106781 | 7.77817459 | 10.6066017 | 7.77817459 | 4.94974747 |
| ENSG00000176597 | B3GNT5 | 2.34146809 | 0.00871728 | 63.6396103 | 57.9827561 | 93.3380951 | 100.409163 | 33.9411255 | 2.12132034 | 24.0416306 | 2.12132034 |
| ENSG00000150782 | IL18 | 2.28950555 | 0.00874201 | 5.65685425 | 12.7279221 | 5.65685425 | 7.07106781 | 1.41421356 | 1.41421356 | 1.41421356 | 2.12132034 |
| ENSG00000224127 | RP11-510C10.2 | 3.16992184 | 0.0088554 | 1.41421356 | 1.41421356 | 24.0416306 | 24.0416306 | 1.41421356 | 1.41421356 | 1.41421356 | 1.41421356 |
| ENSG00000179455 | MKRN3 | 2.67807041 | 0.00907236 | 31.1126984 | 28.2842713 | 2.82842713 | 5.65685425 | 2.12132034 | 1.41421356 | 5.65685425 | 1.41421356 |
| ENSG00000090661 | CERS4 | -2.90688664 | 0.00944549 | 1.41421356 | 1.41421356 | 1.41421356 | 4.24264069 | 4.94974747 | 25.4558441 | 4.24264069 | 28.991378 |
| ENSG00000048540 | LMO3 | 2.13750269 | 0.00958551 | 9.89949494 | 5.65685425 | 7.07106781 | 8.48528137 | 1.41421356 | 1.41421356 | 2.12132034 | 2.12132034 |
| ENSG00000265489 | RP11-1090M7.1 | 2.60768138 | 0.0096019 | 2.82842713 | 4.24264069 | 36.7695526 | 46.6690476 | 3.53553391 | 4.94974747 | 3.53553391 | 2.82842713 |
| ENSG00000228221 | LINC00578 | 3.33038723 | 0.00970021 | 1.41421356 | 2.82842713 | 111.722871 | 125.865007 | 2.12132034 | 7.77817459 | 3.53553391 | 10.6066017 |
| ENSG00000241104 | CEACAMP10 | 2.75488541 | 0.00984416 | 12.7279221 | 19.7989899 | 4.24264069 | 1.41421356 | 1.41421356 | 1.41421356 | 1.41421356 | 1.41421356 |

| DEGs Fibros CHD3 | Symbol | log2FoldChange | padj | p.L915F | p.N1159K | CTL-1 | CTL-2 | CTL-3 | CTL-4 |
|------------------|----------|----------------|------------|---------|----------|-------|-------|-------|-------|
| ENSG00000019186 | CYP24A1 | 3.65207478 | 0.00685758 | 104 | 28 | 1 | 11 | 8 | 1 |
| ENSG00000064042 | LIMCH1 | -4.39430926 | 1.04E-06 | 97 | 75 | 4107 | 647 | 1971 | 509 |
| ENSG00000067715 | SYT1 | -2.88565296 | 0.00018455 | 37 | 36 | 447 | 241 | 266 | 125 |
| ENSG00000075275 | CELSR1 | -3.49911756 | 0.00280458 | 19 | 25 | 480 | 294 | 196 | 25 |
| ENSG00000079841 | RIMS1 | -5.5302628 | 9.02E-09 | 5 | 2 | 236 | 76 | 113 | 222 |
| ENSG00000099204 | ABLM1 | -2.74679381 | 0.00011479 | 155 | 90 | 399 | 1036 | 1087 | 767 |
| ENSG00000101638 | ST8SIA5 | -3.27798246 | 0.00125552 | 5 | 10 | 57 | 48 | 134 | 52 |
| ENSG00000106511 | MEOX2 | 6.09880018 | 1.32E-14 | 258 | 1147 | 5 | 22 | 8 | 6 |
| ENSG00000106565 | TMEM176B | 4.02974486 | 0.00200141 | 82 | 16 | 4 | 2 | 5 | 1 |
| ENSG00000106688 | SLC1A1 | -3.25169943 | 0.00945704 | 156 | 85 | 1453 | 203 | 2616 | 319 |
| ENSG00000106772 | PRUNE2 | -2.68799066 | 0.00244924 | 326 | 462 | 4635 | 802 | 2580 | 2139 |
| ENSG00000110203 | FOLR3 | -4.15152403 | 0.00018455 | 22 | 35 | 1130 | 45 | 503 | 348 |
| ENSG00000115380 | EFEMP1 | 2.66616771 | 0.00143314 | 22576 | 15610 | 1329 | 5257 | 1718 | 3728 |
| ENSG00000115604 | IL18R1 | 4.69031192 | 0.00055442 | 133 | 9 | 3 | 1 | 5 | 2 |
| ENSG00000118785 | SPP1 | -3.29191253 | 0.00361199 | 22 | 12 | 334 | 29 | 162 | 141 |
| ENSG00000118898 | PPL | 3.0472351 | 0.00044792 | 1470 | 1262 | 173 | 260 | 192 | 36 |
| ENSG00000126217 | MCF2L | -3.78509777 | 0.00260448 | 16 | 26 | 382 | 26 | 122 | 628 |
| ENSG00000128655 | PDE11A | -4.98013427 | 4.59E-05 | 6 | 2 | 155 | 13 | 171 | 166 |
| ENSG00000130600 | H19 | 4.8324995 | 3.27E-13 | 637 | 289 | 11 | 14 | 30 | 10 |
| ENSG00000135324 | MRAP2 | -4.18914053 | 3.13E-08 | 13 | 16 | 310 | 161 | 451 | 136 |
| ENSG00000137203 | TFAP2A | -4.35434546 | 1.16E-05 | 34 | 32 | 655 | 1558 | 98 | 389 |
| ENSG00000144891 | AGTR1 | 2.50875187 | 0.00462124 | 385 | 887 | 53 | 77 | 184 | 133 |
| ENSG00000147655 | RSPO2 | 4.29767657 | 0.00705694 | 13 | 46 | 1 | 3 | 1 | 1 |
| ENSG00000148426 | PROSER2 | -2.68011779 | 0.00720879 | 33 | 22 | 271 | 82 | 268 | 84 |
| ENSG00000149968 | MMP3 | 2.65190081 | 0.00351447 | 13257 | 40495 | 5274 | 1532 | 4128 | 6171 |
| ENSG00000152268 | SPON1 | 3.30527634 | 2.49E-05 | 443 | 160 | 21 | 17 | 48 | 36 |
| ENSG00000152578 | GRIA4 | 4.74415677 | 0.00078884 | 23 | 44 | 1 | 1 | 2 | 1 |
| ENSG00000152580 | IGSF10 | 3.90175881 | 7.05E-08 | 750 | 513 | 31 | 40 | 82 | 16 |

| | | | | | | | | | |
|------------------------|-----------|-------------|------------|------|------|-------|------|-------|-------|
| ENSG00000153002 | CPB1 | 3.887523 | 0.00201661 | 28 | 46 | 2 | 2 | 5 | 1 |
| ENSG00000153132 | CLGN | -2.26429982 | 0.00534491 | 203 | 129 | 798 | 574 | 1350 | 468 |
| ENSG00000156463 | SH3RF2 | -2.07778571 | 0.00941381 | 48 | 40 | 221 | 252 | 133 | 137 |
| ENSG00000156475 | PPP2R2B | -5.73639515 | 2.95E-06 | 5 | 3 | 582 | 30 | 118 | 123 |
| ENSG00000157554 | ERG | -3.44421524 | 0.00742831 | 11 | 2 | 88 | 27 | 134 | 34 |
| ENSG00000158528 | PPP1R9A | -3.91374011 | 0.00534491 | 6 | 1 | 19 | 90 | 78 | 24 |
| ENSG00000162631 | NTNG1 | -2.66481731 | 0.00534491 | 367 | 349 | 782 | 1258 | 4528 | 2513 |
| ENSG00000162692 | VCAM1 | 4.56585075 | 0.00037393 | 216 | 9 | 6 | 5 | 6 | 2 |
| ENSG00000163395 | IGFN1 | 4.47804413 | 0.00042884 | 27 | 51 | 1 | 3 | 2 | 1 |
| ENSG00000164106 | SCRG1 | 4.47186684 | 3.72E-06 | 173 | 60 | 12 | 1 | 4 | 4 |
| ENSG00000164176 | EDIL3 | -4.21354732 | 5.32E-05 | 130 | 117 | 601 | 1604 | 629 | 6331 |
| ENSG00000165105 | RASEF | -3.02333068 | 0.00067059 | 5 | 18 | 70 | 99 | 131 | 74 |
| ENSG00000170893 | TRH | 4.66296151 | 0.00022768 | 35 | 41 | 2 | 1 | 1 | 2 |
| ENSG00000177570 | SAMD12 | -2.09953466 | 0.00303452 | 63 | 77 | 391 | 220 | 354 | 235 |
| ENSG00000178038 | ALS2CL | -2.37696809 | 0.00173525 | 70 | 128 | 441 | 371 | 845 | 400 |
| ENSG00000178184 | PARD6G | -2.29205188 | 0.00353198 | 66 | 217 | 614 | 833 | 587 | 738 |
| ENSG00000180053 | NKX2-6 | -4.22238856 | 0.00056842 | 2 | 4 | 81 | 20 | 49 | 74 |
| ENSG00000183715 | OPCML | -3.512609 | 0.00017143 | 31 | 15 | 220 | 84 | 206 | 540 |
| ENSG00000184058 | TBX1 | 4.42626339 | 1.90E-06 | 725 | 737 | 21 | 41 | 71 | 3 |
| ENSG00000186684 | CYP27C1 | -3.04803012 | 0.00313268 | 17 | 55 | 349 | 166 | 106 | 570 |
| ENSG00000188393 | CLEC2A | 5.48112283 | 1.90E-06 | 30 | 171 | 1 | 1 | 6 | 1 |
| ENSG00000188581 | KRTAP1-1 | -3.63807043 | 0.00118027 | 8 | 22 | 251 | 33 | 117 | 346 |
| ENSG00000196581 | AJAP1 | 5.74953229 | 1.72E-14 | 297 | 241 | 11 | 6 | 1 | 2 |
| ENSG00000197106 | SLC6A17 | -3.22160897 | 0.00056559 | 29 | 70 | 232 | 465 | 955 | 195 |
| ENSG00000197614 | MFAP5 | -2.57373254 | 0.00180091 | 3803 | 1991 | 21431 | 5989 | 19043 | 22526 |
| ENSG00000198729 | PPP1R14C | -4.35754753 | 0.00123443 | 3 | 2 | 100 | 26 | 52 | 27 |
| ENSG00000198848 | CES1 | 4.32527727 | 0.00016264 | 227 | 204 | 1 | 27 | 14 | 1 |
| ENSG00000231389 | HLA-DPA1 | -4.2910782 | 0.00012334 | 7 | 6 | 48 | 84 | 317 | 60 |
| ENSG00000232725 | U52111.14 | 3.76249908 | 0.00043916 | 41 | 54 | 1 | 4 | 4 | 5 |

| | | | | | | | | | |
|------------------------|-----------|-------------|------------|---|---|----|-----|-----|----|
| ENSG00000249242 | TMEM150C | -3.36194124 | 0.00143516 | 9 | 7 | 47 | 121 | 128 | 33 |
| ENSG00000250337 | LINC01021 | -4.13270987 | 0.00055442 | 8 | 4 | 96 | 16 | 211 | 98 |

| DEGs LCLs TRRAP | Symbol | log2FoldChange | padj | p.L805F | p.A1043T_1 | p.A1043T_2 | CTL-1 | CTL-2 | CTL-3 |
|-----------------|-------------|----------------|----------|---------|------------|------------|-------|-------|-------|
| ENSG00000205786 | AC002511.1 | -3.11334 | 0.002155 | 67 | 15 | 100 | 174 | 386 | 1015 |
| ENSG00000170425 | ADORA2B | 2.766366 | 0.002155 | 104 | 96 | 147 | 29 | 5 | 17 |
| ENSG00000272620 | AFAP1-AS1 | -5.09493 | 3.49E-08 | 3 | 6 | 8 | 52 | 393 | 136 |
| ENSG00000214944 | ARHGEF28 | -2.91451 | 0.004531 | 53 | 159 | 27 | 211 | 462 | 1129 |
| ENSG00000169136 | ATF5 | -2.15265 | 0.004531 | 3918 | 2338 | 1877 | 5860 | 13377 | 16926 |
| ENSG00000187479 | C11orf96 | -3.16619 | 1.11E-05 | 24 | 39 | 23 | 121 | 248 | 403 |
| ENSG00000198216 | CACNA1E | -3.02687 | 3.49E-08 | 452 | 225 | 287 | 1873 | 3377 | 2607 |
| ENSG00000205476 | CCDC85C | -2.94385 | 0.005947 | 54 | 35 | 42 | 67 | 176 | 765 |
| ENSG00000129277 | CCL4 | -2.88459 | 0.00463 | 918 | 624 | 1645 | 1987 | 3810 | 17739 |
| ENSG00000163823 | CCR1 | 2.046204 | 0.009093 | 480 | 778 | 836 | 199 | 231 | 77 |
| ENSG00000112486 | CCR6 | 2.287577 | 0.008606 | 118 | 142 | 72 | 34 | 13 | 21 |
| ENSG00000133048 | CHI3L1 | -4.57399 | 1.96E-05 | 1 | 11 | 21 | 37 | 295 | 454 |
| ENSG00000121898 | CPXM2 | -4.22239 | 3.64E-05 | 1 | 13 | 10 | 57 | 78 | 313 |
| ENSG00000147408 | CSGALNACT1 | 3.950611 | 3.00E-05 | 36 | 101 | 64 | 4 | 4 | 5 |
| ENSG00000227117 | CTA-85E5.10 | 2.541146 | 0.009567 | 44 | 122 | 61 | 10 | 19 | 10 |
| ENSG00000263264 | CTB-133G6.1 | 2.258734 | 0.007147 | 180 | 276 | 214 | 61 | 17 | 62 |
| ENSG00000073737 | DHR9 | 3.25298 | 0.007148 | 84 | 20 | 39 | 8 | 4 | 3 |
| ENSG00000138080 | EMILIN1 | 3.417991 | 0.00023 | 834 | 1813 | 3221 | 405 | 87 | 57 |
| ENSG00000124491 | F13A1 | 2.541238 | 0.000157 | 3634 | 2512 | 4838 | 297 | 845 | 745 |
| ENSG00000086570 | FAT2 | -3.31768 | 3.91E-05 | 10 | 10 | 14 | 60 | 91 | 188 |
| ENSG00000134775 | FHOD3 | -3.46566 | 0.002511 | 23 | 218 | 53 | 170 | 1778 | 1300 |
| ENSG00000128573 | FOXP2 | -3.23619 | 0.008543 | 4 | 9 | 13 | 14 | 158 | 73 |
| ENSG00000170324 | FRMPD2 | -2.22181 | 0.007371 | 56 | 46 | 77 | 232 | 138 | 465 |
| ENSG00000137726 | FXYD6 | -4.29767 | 0.004134 | 6 | 2 | 1 | 5 | 49 | 123 |
| ENSG00000110328 | GALNT18 | -2.04175 | 0.004531 | 133 | 90 | 75 | 262 | 560 | 405 |
| ENSG00000164161 | HHIP | 3.30666 | 0.009555 | 21 | 128 | 39 | 5 | 12 | 2 |
| ENSG00000248890 | HHIP-AS1 | 3.607329 | 0.007371 | 18 | 152 | 25 | 5 | 9 | 2 |

| | | | | | | | | | |
|------------------------|---------------|----------|----------|------|------|-----|-----|------|------|
| ENSG00000271361 | HTATSF1P2 | 2.576139 | 0.006119 | 154 | 133 | 202 | 56 | 11 | 15 |
| ENSG00000106302 | HYAL4 | -3.46379 | 0.002434 | 29 | 29 | 2 | 429 | 172 | 61 |
| ENSG00000125538 | IL1B | -2.97304 | 0.001448 | 190 | 142 | 269 | 663 | 548 | 3508 |
| ENSG00000132321 | IQCA1 | 3.374123 | 0.002511 | 20 | 60 | 117 | 9 | 6 | 4 |
| ENSG00000163207 | IVL | -4.30117 | 4.36E-06 | 9 | 24 | 9 | 136 | 615 | 77 |
| ENSG00000011201 | KAL1 | 4.014948 | 0.007372 | 4 | 36 | 154 | 3 | 6 | 3 |
| ENSG00000145936 | KCNMB1 | 3.800147 | 0.001227 | 54 | 934 | 210 | 60 | 12 | 14 |
| ENSG00000144893 | MED12L | 2.877447 | 0.009555 | 164 | 37 | 284 | 41 | 11 | 14 |
| ENSG00000178053 | MLF1 | -3.04525 | 0.002971 | 18 | 22 | 11 | 36 | 107 | 278 |
| ENSG00000008441 | NFIX | -2.05355 | 0.004128 | 37 | 39 | 43 | 138 | 139 | 217 |
| ENSG00000074527 | NTN4 | 4.149745 | 0.002511 | 25 | 247 | 12 | 6 | 3 | 7 |
| ENSG00000108405 | P2RX1 | 2.419376 | 0.001227 | 146 | 141 | 157 | 43 | 23 | 17 |
| ENSG00000188389 | PDCD1 | -2.83903 | 0.004531 | 208 | 72 | 119 | 299 | 486 | 2070 |
| ENSG00000162896 | PIGR | -2.74115 | 0.005998 | 480 | 168 | 505 | 548 | 4759 | 2402 |
| ENSG00000254681 | PKD1P5 | 3.3025 | 0.004134 | 41 | 1152 | 356 | 55 | 42 | 60 |
| ENSG00000007038 | PRSS21 | -3.58775 | 1.96E-05 | 8 | 22 | 13 | 271 | 58 | 188 |
| ENSG00000144724 | PTPRG | -3.04062 | 0.004546 | 174 | 19 | 74 | 257 | 1415 | 525 |
| ENSG00000154917 | RAB6B | -4.44294 | 0.001227 | 5 | 4 | 3 | 10 | 27 | 224 |
| ENSG00000138670 | RASGEF1B | 3.0614 | 0.004531 | 78 | 33 | 81 | 13 | 8 | 2 |
| ENSG00000058335 | RASGRF1 | -2.6115 | 0.001679 | 166 | 128 | 101 | 415 | 417 | 1582 |
| ENSG00000254006 | RP11-1D12.2 | -2.96657 | 0.000961 | 40 | 11 | 80 | 357 | 192 | 475 |
| ENSG00000260303 | RP11-203B7.2 | 2.365396 | 0.004888 | 222 | 131 | 85 | 36 | 17 | 32 |
| ENSG00000260896 | RP11-314O13.1 | -4.33955 | 3.49E-08 | 12 | 9 | 40 | 644 | 157 | 434 |
| ENSG00000230880 | RP11-417J8.3 | 5.294615 | 0.001483 | 37 | 4 | 116 | 1 | 2 | 1 |
| ENSG00000182568 | SATB1 | 3.263034 | 0.00938 | 396 | 58 | 26 | 23 | 18 | 9 |
| ENSG00000074660 | SCARF1 | 2.584308 | 0.000201 | 1141 | 2396 | 871 | 357 | 185 | 193 |
| ENSG00000151967 | SCHIP1 | 3.867895 | 0.001043 | 15 | 87 | 263 | 13 | 7 | 5 |
| ENSG00000135094 | SDS | -4.97345 | 1.42E-06 | 5 | 4 | 3 | 29 | 84 | 264 |
| ENSG00000070915 | SLC12A3 | -3.32192 | 0.007147 | 13 | 3 | 5 | 36 | 148 | 26 |

| | | | | | | | | | |
|------------------------|--------|----------|----------|-----|------|-----|----|-----|----|
| ENSG00000186212 | SOWAHB | -3.93411 | 0.006805 | 2 | 2 | 3 | 9 | 39 | 59 |
| ENSG00000077327 | SPAG6 | 3.377554 | 0.000255 | 82 | 442 | 983 | 35 | 58 | 52 |
| ENSG00000164691 | TAGAP | 2.265402 | 0.007187 | 260 | 569 | 147 | 65 | 55 | 83 |
| ENSG00000041982 | TNC | -3.02106 | 3.55E-05 | 10 | 12 | 12 | 86 | 112 | 78 |
| ENSG00000115085 | ZAP70 | 4.366811 | 1.57E-07 | 436 | 1554 | 589 | 79 | 38 | 8 |
| ENSG00000183850 | ZNF730 | 2.593326 | 6.03E-05 | 153 | 229 | 137 | 29 | 33 | 24 |
| ENSG00000124203 | ZNF831 | 2.542898 | 0.002511 | 85 | 202 | 220 | 46 | 18 | 23 |

Annexe II : Pics d'ATACseq significatifs pour les fibroblastes *TRRAP* p.W1866C

| PeakID | Chr | Start | End | Strand | Annotation | Distance to TSS | Nearest PromoterID | Gene Name |
|-------------------|-------|----------|----------|--------|--------------------------------------|-----------------|--------------------|------------|
| chr13-4884 | chr13 | 1.12E+08 | 1.12E+08 | + | Intergenic | 78340 | NM_152324 | TEX29 |
| chr22-259 | chr22 | 50118402 | 50119152 | + | Intergenic | -67587 | NR_110522 | C22orf34 |
| chr8-14649 | chr8 | 40470921 | 40471685 | + | intron (NM_001135731, intron 4 of 5) | 284040 | NM_001135731 | ZMAT4 |
| chr2-23207 | chr2 | 1.05E+08 | 1.05E+08 | + | intron (NR_038231, intron 1 of 3) | 2221 | NR_038231 | LINC01114 |
| chr6-26983 | chr6 | 57538282 | 57541018 | + | Intergenic | 284720 | NR_036097 | MIR548U |
| chr13-1312 | chr13 | 98758081 | 98758706 | + | Intergenic | -36423 | NM_001286839 | FARP1 |
| chr7-23351 | chr7 | 1.4E+08 | 1.4E+08 | + | intron (NM_004333, intron 17 of 17) | -41431 | NR_024454 | NDUFB2-AS1 |

| | | | | | | | | |
|--------------------------------|--------------------------|----------|----------|---|--|---------|--------------|-----------|
| chr22-4308 | chr22 | 25463595 | 25464744 | + | intron (NM_001145206, intron 4 of 10) | 40228 | NM_001145206 | KIAA1671 |
| chr19-1246 | chr19 | 15562985 | 15563865 | + | intron (NM_022904, intron 16 of 17) | -2663 | NM_021241 | WIZ |
| chr2-5826 | chr2 | 1.55E+08 | 1.55E+08 | + | Intergenic | -74374 | NM_002239 | KCNJ3 |
| chr8-15135 | chr8 | 87979181 | 87980328 | + | intron (NM_173538, intron 4 of 10) | 101078 | NM_173538 | CNBD1 |
| chr16-18139 | chr16 | 77517340 | 77518532 | + | Intergenic | -48925 | NM_199355 | ADAMTS18 |
| chr6-19061 | chr6 | 50348064 | 50349236 | + | Intergenic | -332286 | NM_001037498 | DEFB112 |
| chr1-8852 | chr1 | 74077357 | 74077983 | + | Intergenic | 305817 | NR_110676 | LINC01360 |
| chr11-20041 | chr11 | 65198854 | 65205969 | + | non-coding (NR_131012, exon 1 of 1) | -9518 | NR_030343 | MIR612 |
| chr4-17939 | chr4 | 15119631 | 15120774 | + | intron (NR_125911, intron 2 of 5) | 115904 | NM_001177383 | CPEB2 |
| chr4_gl000193_random-13 | chr4_ gl000193_random | 83007 | 87406 | + | intron (NR_038377, intron 1 of 5) | 3169 | NR_038377 | LINC01667 |
| chr6-23892 | chr6 | 88653891 | 88654880 | + | Intergenic | -103122 | NM_030960 | SPACA1 |
| chr7-14919 | chr7 | 1.59E+08 | 1.59E+08 | + | intron (NR_130758, intron 4 of 12) | -1979 | NM_001308259 | VIPR2 |
| chr5-8827 | chr5 | 1.14E+08 | 1.14E+08 | + | Intergenic | 135919 | NM_001300759 | TRIM36 |
| chrX-12878 | chrX | 79540609 | 79541430 | + | intron (NR_110646, intron 5 of 5) | 49798 | NR_110646 | CHMP1B2P |
| chr2-41746 | chr2 | 1.51E+08 | 1.51E+08 | + | Intergenic | 22572 | NM_005168 | RND3 |
| chrUn_gl000219-40 | chrUn n_gl000219 | 132667 | 133673 | + | Intergenic | -33528 | NR_027436 | LOC283788 |

| | | | | | | | | |
|--------------------------|-----------------------------|----------|----------|---|---|---------|------------------|--------------|
| chr13-6623 | chr13 | 36488474 | 36489964 | + | intron (NM_001330071, intron 4 of 16) | -59221 | NM_0011954 15 | DCLK1 |
| chrUn_gl000224-42 | chrUn n_gl0 0022 4 | 212 | 3307 | + | Intergenic | -100744 | NR_135482 | LOC100505874 |
| chr11-3528 | chr11 | 73729444 | 73730199 | + | intron (NM_001286577, intron 32 of 32) | -9539 | NM_003356 | UCP3 |
| chr8-20672 | chr8 | 43252013 | 43253177 | + | Intergenic | 105010 | NM_0010053 65 | POTEA |
| chr11-21271 | chr11 | 61728612 | 61732063 | + | non-coding (NR_134580, exon 10 of 11) | 4795 | NM_002032 | FTH1 |
| chr9-19240 | chr9 | 1.21E+08 | 1.21E+08 | + | Intergenic | 829095 | NM_138554 | TLR4 |
| chr21-1228 | chr21 | 44573611 | 44574236 | + | Intergenic | -12721 | NR_133678 | LOC106780825 |
| chr9-20347 | chr9 | 1.09E+08 | 1.09E+08 | + | Intergenic | 163725 | NM_018112 | TMEM38B |
| chr13-2785 | chr13 | 1.02E+08 | 1.02E+08 | + | intron (NR_047687, intron 5 of 7) | -105361 | NR_047015 | LINC00411 |
| chr1-33530 | chr1 | 41471438 | 41473676 | + | intron (NR_125440, intron 13 of 17) | -7705 | NR_037868 | SLFN1-AS1 |
| chr19-6942 | chr19 | 55700540 | 55701443 | + | intron (NM_001161440, intron 10 of 17) | -9271 | NM_003180 | SYT5 |
| chr2-39106 | chr2 | 90007568 | 90008500 | + | Intergenic | 896150 | NR_039635 | MIR4436A |
| chr14-256 | chr14 | 37130862 | 37133579 | + | exon (NM_006194, exon 3 of 5) | 5447 | NM_006194 | PAX9 |
| chr5-26320 | chr5 | 10607481 | 10608520 | + | intron (NM_001164440, intron 1 of 3) | 43565 | NM_0011644 40 | ANKRD33B |
| chr5-7677 | chr5 | 21341178 | 21342449 | + | Intergenic | -117776 | NR_027027 | GUSBP1 |
| chr6-29579 | chr6 | 10659419 | 10660852 | + | Intergenic | 34704 | NM_0011450 20 | C6orf52 |
| chr9-136 | chr9 | 69710795 | 69711420 | + | intron (NR_135597, intron 4 of 4).2 | 46677 | NR_121570 | LOC101928381 |
| chr20-10936 | chr20 | 21434063 | 21435121 | + | Intergenic | -56545 | NM_033176 | NKX2-4 |

| | | | | | | | | |
|--------------------------|-----------------------------|----------|----------|---|---|---------|------------------|--------------|
| chr6-17209 | chr6 | 76095756 | 76096665 | + | intron (NR_125859, intron 2 of 8) | 3308 | NR_125859 | LOC101928540 |
| chrUn_gl000224-18 | chrUn n_gl0 0022 4 | 6478 | 8690 | + | Intergenic | -94919 | NR_135482 | LOC100505874 |
| chrX-3652 | chrX | 1.23E+08 | 1.23E+08 | + | Intergenic | -12954 | NM_0012044 01 | XIAP |
| chr13-10937 | chr13 | 41786780 | 41787889 | + | intron (NR_120423, intron 3 of 3) | -18623 | NM_032138 | KBTBD7 |
| chr12-9635 | chr12 | 77024160 | 77024961 | + | Intergenic | -70971 | NM_0013196 52 | OSBPL8 |
| chr1-32340 | chr1 | 1.43E+08 | 1.43E+08 | + | Intergenic | -174074 | NR_106967 | MIR6077 |
| chr1-49474 | chr1 | 1.75E+08 | 1.75E+08 | + | TTS (NM_001243764) | -3279 | NM_0010072 14 | CACYBP |
| chr4-22083 | chr4 | 1.7E+08 | 1.7E+08 | + | Intergenic | -64427 | NM_020870 | SH3RF1 |
| chr8-22051 | chr8 | 41258386 | 41259877 | + | Intergenic | -88950 | NM_016099 | GOLGA7 |
| chr18-7675 | chr18 | 14415576 | 14416328 | + | Intergenic | 78530 | NR_026756 | CYP4F35P |
| chr2-29289 | chr2 | 90012275 | 90013284 | + | Intergenic | 900895 | NR_039635 | MIR4436A |
| chr4-7844 | chr4 | 68099805 | 68100560 | + | Intergenic | 187536 | NR_110747 | LOC101927237 |
| chr2-11101 | chr2 | 2.25E+08 | 2.25E+08 | + | Intergenic | 15338 | NM_022915 | MRPL44 |
| chr20-2971 | chr20 | 4080478 | 4081105 | + | Intergenic | -48635 | NM_175840 | SMOX |
| chr5-3265 | chr5 | 71623886 | 71624511 | + | intron (NM_001284404, intron 3 of 7) | 8004 | NM_0012844 04 | PTCD2 |
| chr16-11816 | chr16 | 65234740 | 65236004 | + | Intergenic | -24896 | NR_110918 | LINC02126 |
| chr6-11624 | chr6 | 47128718 | 47129473 | + | Intergenic | -118996 | NM_153840 | ADGRF1 |
| chr14-11144 | chr14 | 1.05E+08 | 1.05E+08 | + | Intergenic | 10999 | NM_0013081 74 | PLD4 |
| chr20-10191 | chr20 | 30084292 | 30086156 | + | Intergenic | 11643 | NR_024358 | LINC00028 |
| chr2-32290 | chr2 | 1.47E+08 | 1.47E+08 | + | Intergenic | -825955 | NR_026904 | PABPC1P2 |

| | | | | | | | | |
|---------------------------|-----------------------------|----------|----------|---|--|---------|------------------|--------------------------------|
| chr3-2451 | chr3 | 1.11E+08 | 1.11E+08 | + | intron (NM_005816, intron 8 of 13) | -21496 | NM_024508 | ZBED2 |
| chr19-12676 | chr19 | 53811446 | 53812212 | + | Intergenic | -16954 | NM_033341 | BIRC8 |
| chr4-25906 | chr4 | 32500524 | 32501370 | + | Intergenic | 148287 | NR_134672 | LOC101927363 |
| chr15-14885 | chr15 | 34034365 | 34035699 | + | intron (NM_001243996, intron 52 of 102) | -225483 | NM_0013209 17 | CHRM5 |
| chr3-5201 | chr3 | 1.01E+08 | 1.01E+08 | + | Intergenic | 103607 | NM_016247 | IMPG2 |
| chr6-18620 | chr6 | 1.41E+08 | 1.41E+08 | + | Intergenic | 296413 | NR_039675 | MIR4465 |
| chrUn_gli000224-70 | chrU n_gli0 0022 4 | 17134 | 20991 | + | Intergenic | -83441 | NR_135482 | LOC100505874 |
| chr15-685 | chr15 | 39873531 | 39892314 | + | intron (NM_003246, intron 14 of 21) | 9642 | NM_003246 | THBS1 |
| chr3-4440 | chr3 | 7042172 | 7042797 | + | intron (NM_000844, intron 1 of 9) | -106780 | NR_131906 | GRM7-AS2 |
| chr17-15321 | chr17 | 22203791 | 22204567 | + | Intergenic | 181742 | NM_0011904 52 | MTRNR2L1 |
| chr1-16457 | chr1 | 6698064 | 6698836 | + | intron (NM_018198, intron 11 of 15) | 13240 | NM_0011957 52 | THAP3 |
| chr21-4073 | chr21 | 15352097 | 15353014 | + | non-coding (NR_027270, exon 1 of 6) | 210 | NR_027270 | ANKRD20A11P |
| chr6-29197 | chr6 | 58385554 | 58386536 | + | Intergenic | -98321 | NR_125729 | LINC00680 |
| chr7-2789 | chr7 | 76637297 | 76637922 | + | intron (NR_023383, intron 6 of 10) | 27470 | NR_023383 | DTX2P1- UPK3BP1- PMS2P11 |
| chr2-42878 | chr2 | 2.16E+08 | 2.16E+08 | + | intron (NM_054034, intron 6 of 12) | 10637 | NM_212482 | FN1 |
| chr4-4782 | chr4 | 35910132 | 35910883 | + | Intergenic | -335231 | NR_122079 | LOC439933 |
| chr22-5369 | chr22 | 18574279 | 18575496 | + | Intergenic | 14127 | NM_0011276 49 | PEX26 |
| chr5-14966 | chr5 | 33032675 | 33033553 | + | Intergenic | 85565 | NR_033832 | LINC02120 |

| | | | | | | | | |
|--------------------------|----------------|----------|----------|---|--|---------|------------------|--------------|
| chr17-6804 | chr17 | 19630535 | 19632141 | + | Intergenic | -9046 | NR_135624 | SLC47A2 |
| chrX-3180 | chrX | 79471586 | 79472338 | + | Intergenic | 118855 | NR_110646 | CHMP1B2P |
| chr6-15535 | chr6 | 69814975 | 69815754 | + | intron (NM_001704, intron 17 of 31) | 469732 | NM_001704 | ADGRB3 |
| chr14-8155 | chr14 | 52830428 | 52831212 | + | Intergenic | 49804 | NM_000956 | PTGER2 |
| chr14-2731 | chr14 | 20291858 | 20292483 | + | intron (NM_001004723, intron 1 of 1) | 20241 | NM_0010047 23 | OR4N2 |
| chrX-4661 | chrX | 57633889 | 57634783 | + | Intergenic | 16067 | NM_007157 | ZXDB |
| chr19-10019 | chr19 | 57813035 | 57814085 | + | Intergenic | -18305 | NM_213598 | ZNF543 |
| chr13-6148 | chr13 | 28427813 | 28428726 | + | intron (NR_047484, intron 1 of 3) | 61489 | NM_145657 | GSX1 |
| chr22-5998 | chr22 | 35931792 | 35932879 | + | Intergenic | -5017 | NM_014310 | RASD2 |
| chr2-29939 | chr2 | 2.25E+08 | 2.25E+08 | + | intron (NM_006216, intron 1 of 8) | 26878 | NR_073116 | SERPINE2 |
| chr18-7448 | chr18 | 14969791 | 14972538 | + | promoter-TSS (NR_110783) | -697 | NR_110783 | LINC01444 |
| chr14-7514 | chr14 | 33583113 | 33584006 | + | intron (NM_001164749, intron 2 of 11) | 175100 | NM_0011647 49 | NPAS3 |
| chr2-41553 | chr2 | 91749832 | 91751461 | + | Intergenic | 97329 | NR_027238 | LOC654342 |
| chr3-19828 | chr3 | 4813474 | 4814503 | + | intron (NM_002222, intron 42 of 57) | -20714 | NR_004428 | EGOT |
| chr1-41015 | chr1 | 1.53E+08 | 1.53E+08 | + | Intergenic | -5857 | NM_0011286 00 | LCE6A |
| chr5-11940 | chr5 | 17061988 | 17062881 | + | Intergenic | -126049 | NM_012334 | MYO10 |
| chr3-24262 | chr3 | 1.84E+08 | 1.84E+08 | + | intron (NM_001322103, intron 3 of 3) | 27706 | NM_0013221 03 | LINC02054 |
| chr12-17832 | chr12 | 1.2E+08 | 1.2E+08 | + | intron (NM_001136534, intron 1 of 2) | 1444 | NM_0011365 34 | TMEM233 |
| chr4-20453 | chr4 | 31540701 | 31541813 | + | Intergenic | 368491 | NR_125935 | LOC102723778 |
| chrUn_gl000224-34 | chrUn n_gl0 | 11024 | 11927 | + | Intergenic | -91028 | NR_135482 | LOC100505874 |

| | | | | | | | | |
|--------------------|-------|----------|----------|---|--------------------------------------|---------|--------------|--------------|
| | 0022 | | | | | | | |
| | 4 | | | | | | | |
| chr11-11418 | chr11 | 59230483 | 59231387 | + | Intergenic | 6501 | NM_001004708 | OR4D6 |
| chr14-15182 | chr14 | 24109051 | 24110616 | + | intron (NM_005794, intron 4 of 8) | 4260 | NM_182908 | DHRS2 |
| chr15-3346 | chr15 | 86790760 | 86791386 | + | exon (NM_152336, exon 7 of 25) | 69126 | NR_046012 | AGBL1-AS1 |
| chr12-15589 | chr12 | 86036377 | 86037168 | + | Intergenic | 193546 | NM_005447 | RASSF9 |
| chr4-8566 | chr4 | 35794730 | 35795747 | + | Intergenic | -450500 | NR_122079 | LOC439933 |
| chr18-4812 | chr18 | 48185217 | 48186122 | + | intron (NM_001292039, intron 1 of 4) | 99185 | NM_001292040 | MAPK4 |
| chr2-33993 | chr2 | 1.56E+08 | 1.56E+08 | + | Intergenic | 552389 | NM_001260510 | KCNJ3 |
| chr10-17781 | chr10 | 1.11E+08 | 1.11E+08 | + | Intergenic | 895290 | NM_001324136 | XPNPEP1 |
| chrX-9823 | chrX | 68436329 | 68437355 | + | Intergenic | 37442 | NR_103715 | LINC00269 |
| chr10-4865 | chr10 | 67118276 | 67118904 | + | Intergenic | -212593 | NR_120647 | LINC01515 |
| chr3-28738 | chr3 | 1.4E+08 | 1.4E+08 | + | intron (NM_022131, intron 9 of 16) | -36968 | NR_108084 | CLSTN2-AS1 |
| chr11-18945 | chr11 | 32872707 | 32874296 | + | intron (NM_024081, intron 5 of 5) | 22020 | NM_024081 | PRRG4 |
| chr5-4491 | chr5 | 52166248 | 52166873 | + | intron (NM_181501, intron 6 of 28) | 82424 | NM_181501 | ITGA1 |
| chr12-16638 | chr12 | 48361313 | 48362488 | + | intron (NM_001143842, intron 7 of 7) | 4570 | NM_001143841 | TMEM106C |
| chr14-11308 | chr14 | 21466392 | 21467821 | + | promoter-TSS (NM_001256588) | 216 | NR_110139 | LOC101929718 |
| chr17-15420 | chr17 | 66165073 | 66166124 | + | Intergenic | -29203 | NR_027283 | LOC440461 |
| chr12-8336 | chr12 | 19555551 | 19556479 | + | Intergenic | -36593 | NM_001114176 | AEBP2 |

| | | | | | | | | |
|--------------------|-------|----------|----------|---|--|----------|------------------|--------------|
| chr18-2905 | chr18 | 13823756 | 13825856 | + | promoter-TSS (NM_005913) | -737 | NM_005913 | MC5R |
| chr6-15463 | chr6 | 13984629 | 13985523 | + | Intergenic | 59652 | NM_0011650 33 | RNF182 |
| chr17-10110 | chr17 | 72565281 | 72566427 | + | Intergenic | -2372 | NM_0013240 76 | LOC100130520 |
| chrX-13440 | chrX | 64461029 | 64461924 | + | Intergenic | -206852 | NM_0011780 32 | ZC4H2 |
| chr18-5796 | chr18 | 72924305 | 72926268 | + | intron (NM_001308210, intron 1 of 1) | 1802 | NM_0013082 10 | TSHZ1 |
| chr5-7301 | chr5 | 45113827 | 45114454 | + | Intergenic | 305113 | NM_016640 | MRPS30 |
| chr2-9060 | chr2 | 1.28E+08 | 1.28E+08 | + | Intergenic | 78380 | NM_0013206 34 | BIN1 |
| chr16-11137 | chr16 | 66398101 | 66399990 | + | Intergenic | -1465 | NM_001795 | CDH5 |
| chr7-13062 | chr7 | 1.04E+08 | 1.04E+08 | + | intron (NM_199000, intron 1 of 2) | 177936 | NR_034142 | LHFPL3-AS1 |
| chr18-6142 | chr18 | 14430509 | 14431977 | + | Intergenic | 67462 | NR_024076 | CXADRP3 |
| chr8-15354 | chr8 | 70091759 | 70092527 | + | Intergenic | -75718 | NR_039986 | LINC01592 |
| chr1-22784 | chr1 | 59495810 | 59496834 | + | intron (NR_110626, intron 2 of 2) | 10174 | NR_110626 | LINC01358 |
| chr1-23978 | chr1 | 1.76E+08 | 1.76E+08 | + | Intergenic | -107107 | NM_0012866 44 | RFWD2 |
| chr18-6672 | chr18 | 7174462 | 7175263 | + | Intergenic | -56275 | NM_0011055 81 | LRRC30 |
| chr13-5224 | chr13 | 43559484 | 43560363 | + | intron (NM_001002264, intron 1 of 12) | 6525 | NM_0010022 64 | EPSTI1 |
| chr14-17162 | chr14 | 83683366 | 83684535 | + | Intergenic | -1594545 | NR_110074 | LINC01467 |
| chr18-6278 | chr18 | 61572468 | 61575430 | + | Intergenic | -1275 | NM_005024 | SERPINB10 |
| chr18-7569 | chr18 | 14469693 | 14471529 | + | Intergenic | 28094 | NR_024076 | CXADRP3 |
| chr1-20873 | chr1 | 2.09E+08 | 2.09E+08 | + | Intergenic | -162463 | NR_135085 | LINC01696 |
| chr9-17816 | chr9 | 68402364 | 68403907 | + | Intergenic | -12173 | NR_039689 | MIR4477B |

| | | | | | | | | |
|--------------------|-------|----------|----------|---|--|---------|------------------|--------------|
| chr7-22350 | chr7 | 89623552 | 89624437 | + | intron (NR_110029, intron 2 of 2) | -124720 | NR_003551 | DPY19L2P4 |
| chr2-36172 | chr2 | 28631586 | 28632900 | + | intron (NM_005253, intron 3 of 3) | -14704 | NR_103831 | FLJ31356 |
| chr2-17998 | chr2 | 1.57E+08 | 1.57E+08 | + | intron (NR_110250, intron 1 of 3) | 139391 | NR_110250 | LINC01876 |
| chr21-3787 | chr21 | 9849932 | 9852942 | + | Intergenic | 25234 | NR_037458 | MIR3687-1 |
| chr1-10682 | chr1 | 2.12E+08 | 2.12E+08 | + | Intergenic | 26833 | NM_021194 | SLC30A1 |
| chr11-9657 | chr11 | 1.32E+08 | 1.32E+08 | + | intron (NM_001144058, intron 3 of 7) | -45666 | NR_046820 | NTM-IT |
| chr13-9103 | chr13 | 95041545 | 95043633 | + | intron (NM_005708, intron 7 of 8) | 89347 | NM_001922 | DCT |
| chr5-3033 | chr5 | 78048922 | 78049547 | + | Intergenic | -104586 | NM_005779 | LHFPL2 |
| chr3-4482 | chr3 | 1.92E+08 | 1.92E+08 | + | Intergenic | -45184 | NM_004113 | FGF12 |
| chr18-2674 | chr18 | 14868766 | 14869266 | + | Intergenic | 38851 | NR_036153 | MIR3156-2 |
| chr15-9942 | chr15 | 59886449 | 59887340 | + | Intergenic | -17088 | NM_004751 | GCNT3 |
| chr7-4231 | chr7 | 1.01E+08 | 1.01E+08 | + | exon (NM_000602, exon 5 of 9) | 6782 | NM_000602 | SERPINE1 |
| chr2-29827 | chr2 | 1.75E+08 | 1.75E+08 | + | Intergenic | -1305 | NM_0011452 50 | SP9 |
| chr16-18006 | chr16 | 46590496 | 46592388 | + | intron (NR_026556, intron 4 of 16) | 11567 | NR_026556 | ANKRD26P1 |
| chr8-16861 | chr8 | 54908294 | 54909186 | + | intron (NR_109901, intron 4 of 11) | 26276 | NR_109902 | TCEA1 |
| chr6-2551 | chr6 | 83420829 | 83422298 | + | Intergenic | 347602 | NM_0011663 92 | TPBG |
| chr5-14766 | chr5 | 1.75E+08 | 1.75E+08 | + | Intergenic | 208769 | NM_000794 | DRD1 |
| chr2-21518 | chr2 | 2247167 | 2248075 | + | intron (NM_001329845, intron 2 of 24) | -75383 | NR_024468 | MYT1L-AS1 |
| chr12-20474 | chr12 | 82281562 | 82282889 | + | Intergenic | -65273 | NR_110089 | LOC101928449 |
| chr21-4771 | chr21 | 34216333 | 34217906 | + | Intergenic | -31066 | NM_0011624 96 | C21orf62 |

| | | | | | | | | |
|--------------------|-------|----------|----------|---|--|---------|--------------|--------------|
| chr10-12079 | chr10 | 36111002 | 36111793 | + | Intergenic | 44167 | NR_110138 | PCAT5 |
| chr2-9488 | chr2 | 90143943 | 90144592 | + | Intergenic | 1032383 | NR_039635 | MIR4436A |
| chr1-15817 | chr1 | 1.53E+08 | 1.53E+08 | + | Intergenic | -6627 | NM_178354 | LCE1F |
| chr14-5255 | chr14 | 20683328 | 20683954 | + | Intergenic | -8228 | NM_001004480 | OR11H6 |
| chr4-20578 | chr4 | 94953747 | 94954534 | + | Intergenic | 174567 | NR_125922 | LOC101929210 |
| chr18-7713 | chr18 | 34073089 | 34073900 | + | intron (NM_025135, intron 3 of 24) | 169887 | NR_132982 | LOC105372071 |
| chr19-13786 | chr19 | 658811 | 660168 | + | intron (NM_194460, intron 1 of 8) | 3744 | NM_194460 | RNF126 |
| chr2-14708 | chr2 | 1.23E+08 | 1.23E+08 | + | Intergenic | 481300 | NM_004622 | TSN |
| chrX-9760 | chrX | 16866450 | 16867228 | + | intron (NM_001198719, intron 10 of 11) | 21139 | NM_001198719 | RBBP7 |
| chr16-12157 | chr16 | 56135853 | 56136912 | + | intron (NR_027078, intron 2 of 4) | 88624 | NR_027078 | LOC283856 |
| chr4-15346 | chr4 | 32931502 | 32932289 | + | Intergenic | 579235 | NR_134672 | LOC101927363 |
| chr2-12541 | chr2 | 2.16E+08 | 2.16E+08 | + | Intergenic | 44903 | NM_004044 | ATIC |
| chr13-2427 | chr13 | 23980998 | 23981624 | + | intron (NM_014363, intron 2 of 9) | -11799 | NR_103450 | SACS-AS1 |
| chr2-45684 | chr2 | 1.55E+08 | 1.55E+08 | + | intron (NM_052917, intron 2 of 12) | -3727 | NM_001301627 | GALNT13 |
| chr12-23022 | chr12 | 1.28E+08 | 1.28E+08 | + | Intergenic | -3914 | NR_120451 | LOC101927637 |
| chr3-26708 | chr3 | 1.66E+08 | 1.66E+08 | + | Intergenic | -66500 | NM_000055 | BCHE |
| chr11-29327 | chr11 | 31408635 | 31409474 | + | intron (NM_181706, intron 2 of 4) | 17677 | NM_181706 | DNAJC24 |
| chr1-1709 | chr1 | 86046841 | 86049511 | + | exon (NM_001554, exon 4 of 5) | 1732 | NM_001554 | CYR61 |
| chr18-3599 | chr18 | 34324069 | 34324945 | + | intron (NM_025135, intron 20 of 24) | -81126 | NR_132982 | LOC105372071 |
| chr18-4499 | chr18 | 19859755 | 19860816 | + | Intergenic | 110887 | NM_005257 | GATA6 |
| chr13-4980 | chr13 | 98064143 | 98065035 | + | Intergenic | -21886 | NM_021033 | RAP2A |

| | | | | | | | | |
|--------------------------|-----------------------------|----------|----------|---|---|---------|------------------|--------------|
| chr2-23809 | chr2 | 90018464 | 90019645 | + | Intergenic | 907170 | NR_039635 | MIR4436A |
| chr6-19736 | chr6 | 1.64E+08 | 1.64E+08 | + | Intergenic | 19307 | NR_134620 | LOC102724152 |
| chr2-5534 | chr2 | 2.07E+08 | 2.07E+08 | + | intron (NM_003812, intron 2 of 25) | 33918 | NM_003812 | ADAM23 |
| chr10-10438 | chr10 | 68081616 | 68082771 | + | intron (NM_001127384, intron 12 of 17) | -603599 | NM_178011 | LRRTM3 |
| chrUn_gl000224-57 | chrUn n_gl0 0022 4 | 12813 | 16077 | + | Intergenic | -88058 | NR_135482 | LOC100505874 |
| chr7-7193 | chr7 | 96672389 | 96673141 | + | Intergenic | -18622 | NM_005221 | DLX5 |
| chr11-8353 | chr11 | 59897922 | 59898422 | + | Intergenic | 42035 | NM_0012569 16 | MS4A2 |
| chr3-8337 | chr3 | 1.47E+08 | 1.47E+08 | + | Intergenic | -319102 | NR_134937 | LINC02032 |
| chr2-23936 | chr2 | 2.01E+08 | 2.01E+08 | + | Intergenic | -172253 | NR_004862 | TYW5 |
| chr13-550 | chr13 | 98279764 | 98280389 | + | Intergenic | 193601 | NM_021033 | RAP2A |
| chr11-29319 | chr11 | 27578588 | 27580361 | + | intron (NR_033315, intron 2 of 6) | -42880 | NR_107054 | MIR8087 |
| chr2-29308 | chr2 | 1.01E+08 | 1.01E+08 | + | Intergenic | -10549 | NR_103730 | LINC01104 |
| chr18-7306 | chr18 | 6993866 | 6995106 | + | intron (NM_005559, intron 34 of 62) | 39810 | NR_126040 | LOC101927188 |
| chrX-10417 | chrX | 1.53E+08 | 1.53E+08 | + | intron (NM_001110792, intron 1 of 2) | -20467 | NR_031757 | MIR718 |
| chrX-11866 | chrX | 1.16E+08 | 1.16E+08 | + | Intergenic | -144246 | NM_0013214 13 | LOC100126447 |
| chr5-22756 | chr5 | 911537 | 913408 | + | intron (NM_004237, intron 10 of 12) | 19503 | NM_0011662 60 | TRIP13 |
| chr9-1661 | chr9 | 85937292 | 85938042 | + | intron (NM_001244959, intron 6 of 14) | 9209 | NM_0012449 61 | FRMD3 |
| chr8-9052 | chr8 | 81838155 | 81839313 | + | Intergenic | -51718 | NM_0010337 23 | ZNF704 |
| chr15-3660 | chr15 | 56776381 | 56777011 | + | Intergenic | -19361 | NM_018365 | MNS1 |

| | | | | | | | | |
|--------------------|-------|----------|----------|---|--|---------|--------------|-----------|
| chr15-7084 | chr15 | 63361385 | 63366215 | + | TTS (NM_001018008) | 23164 | NM_001018008 | TPM1 |
| chr16-4172 | chr16 | 3342506 | 3343257 | + | Intergenic | 9394 | NM_005741 | ZNF263 |
| chr4-23639 | chr4 | 35773535 | 35774488 | + | Intergenic | -471727 | NR_122079 | LOC439933 |
| chr5-28835 | chr5 | 95207897 | 95216160 | + | Intergenic | 24092 | NR_026936 | LINC01554 |
| chr3-17947 | chr3 | 13147145 | 13148942 | + | Intergenic | -33426 | NM_001134382 | IQSEC1 |
| chr1-10085 | chr1 | 1.48E+08 | 1.48E+08 | + | intron (NM_001037501, intron 18 of 19) | -19458 | NR_049810 | MIR5087 |
| chr12-17021 | chr12 | 1.15E+08 | 1.15E+08 | + | exon (NM_005996, exon 1 of 7) | 1266 | NM_005996 | TBX3 |
| chr17-5330 | chr17 | 33337049 | 33337674 | + | non-coding (NR_037713, exon 7 of 7) | 29844 | NM_002311 | LIG3 |
| chr5-14322 | chr5 | 55736872 | 55737788 | + | Intergenic | 40266 | NR_104664 | LINC01948 |
| chr6-24354 | chr6 | 1.18E+08 | 1.18E+08 | + | intron (NM_173674, intron 1 of 14) | 5372 | NM_173674 | DCBLD1 |
| chr7-10961 | chr7 | 1.35E+08 | 1.35E+08 | + | Intergenic | -9580 | NR_144293 | AGBL3 |
| chr9-15672 | chr9 | 71061529 | 71062498 | + | intron (NM_021965, intron 6 of 10) | -89231 | NR_121192 | PGM5-AS1 |
| chr18-9218 | chr18 | 14925001 | 14925920 | + | Intergenic | -20806 | NR_104164 | LINC01443 |
| chr1-11099 | chr1 | 77980752 | 77981768 | + | intron (NM_174858, intron 10 of 13) | 167083 | NM_015534 | ZZZ3 |
| chr14-2635 | chr14 | 76155295 | 76155921 | + | intron (NM_015072, intron 5 of 31) | 28057 | NM_015072 | TTLL5 |
| chr3-19048 | chr3 | 12912257 | 12913173 | + | Intergenic | 13995 | NR_136189 | LINC02022 |
| chr9-3679 | chr9 | 35962067 | 35962962 | + | Intergenic | -4363 | NM_019897 | OR2S2 |
| chr21-1472 | chr21 | 41973971 | 41974602 | + | intron (NM_001389, intron 3 of 32) | 28407 | NR_046774 | DSCAM-IT1 |
| chr9-6230 | chr9 | 67926278 | 67927049 | + | promoter-TSS (NM_032250) | -98 | NM_032250 | ANKRD20A1 |
| chr4-5507 | chr4 | 42675032 | 42675657 | + | Intergenic | -16222 | NM_006095 | ATP8A1 |

| | | | | | | | | |
|--------------------|-------|----------|----------|---|---|---------|------------------|--------------|
| chr7-18725 | chr7 | 1.43E+08 | 1.43E+08 | + | intron (NR_046453, intron 1 of 21) | 2746 | NR_046453 | CLCN1 |
| chr17-9508 | chr17 | 25314709 | 25315605 | + | Intergenic | 101110 | NR_135673 | LOC105371703 |
| chr15-9744 | chr15 | 1.01E+08 | 1.01E+08 | + | promoter-TSS (NR_073405) | -155 | NR_073405 | PRKXP1 |
| chr5-23943 | chr5 | 1.71E+08 | 1.71E+08 | + | Intergenic | 35033 | NR_037474 | MIR3912 |
| chr22-8294 | chr22 | 50157396 | 50158464 | + | Intergenic | 61618 | NM_0013048 09 | BRD1 |
| chr10-10090 | chr10 | 1.19E+08 | 1.19E+08 | + | Intergenic | -16413 | NM_173791 | PDZD8 |
| chr5-18777 | chr5 | 1.02E+08 | 1.02E+08 | + | intron (NM_173488, intron 6 of 13) | 48314 | NM_0012890 04 | SLCO6A1 |
| chr1-52629 | chr1 | 1.52E+08 | 1.52E+08 | + | Intergenic | -27851 | NM_0010099 31 | HRNR |
| chr9-9349 | chr9 | 80744785 | 80745544 | + | Intergenic | -98799 | NM_002072 | GNAQ |
| chr2-22687 | chr2 | 1.49E+08 | 1.49E+08 | + | intron (NM_018328, intron 4 of 14) | -251884 | NM_181742 | ORC4 |
| chr20-12380 | chr20 | 865331 | 866249 | + | exon (NM_001322809, exon 4 of 8) | 31187 | NM_0013228 09 | ANGPT4 |
| chr11-19224 | chr11 | 60267793 | 60268684 | + | intron (NM_001164470, intron 2 of 5) | 7987 | NM_0011644 70 | MS4A12 |
| chr19-13432 | chr19 | 47322931 | 47323715 | + | Intergenic | 10638 | NR_024258 | SNAR-E |
| chr2-37390 | chr2 | 2.25E+08 | 2.25E+08 | + | intron (NM_006216, intron 1 of 8) | 14615 | NR_073116 | SERPINE2 |
| chr3-16857 | chr3 | 1.67E+08 | 1.67E+08 | + | intron (NR_126353, intron 3 of 3) | 14225 | NR_126353 | LINC01327 |
| chr2-17103 | chr2 | 1.05E+08 | 1.05E+08 | + | Intergenic | -17134 | NR_038231 | LINC01114 |
| chr1-51757 | chr1 | 2.1E+08 | 2.1E+08 | + | promoter-TSS (NM_005525) | -128 | NM_005525 | HSD11B1 |
| chr1-15123 | chr1 | 1.48E+08 | 1.48E+08 | + | Intergenic | 56620 | NM_0011016 63 | NBPF11 |
| chr14-7585 | chr14 | 24502175 | 24504312 | + | intron (NR_102687, intron 1 of 3) | -2378 | NR_102689 | DHRS4L1 |

| | | | | | | | | |
|--------------------------------|----------------------|----------|----------|---|---------------------------------------|---------|--------------|--------------|
| chr21-855 | chr21 | 31708844 | 31709721 | + | TTS (NM_001077711) | 730 | NM_001077711 | KRTAP27-1 |
| chr15-7354 | chr15 | 34339471 | 34340367 | + | intron (NM_012125, intron 2 of 2) | -8616 | NM_020371 | AVEN |
| chr6-18284 | chr6 | 36648625 | 36658741 | + | 3' UTR (NM_001220777, exon 3 of 3) | 7196 | NM_001220778 | CDKN1A |
| chr8-6985 | chr8 | 83490568 | 83491324 | + | Intergenic | -333393 | NR_134295 | LOC101927141 |
| chr15-13112 | chr15 | 33012686 | 33017724 | + | intron (NM_013372, intron 1 of 1) | -4139 | NR_109767 | LOC100131315 |
| chr3-14777 | chr3 | 1.67E+08 | 1.67E+08 | + | intron (NR_135545, intron 2 of 2) | 26832 | NR_135545 | LOC105374194 |
| chr9-15542 | chr9 | 1.24E+08 | 1.24E+08 | + | intron (NM_032552, intron 1 of 16) | 69520 | NM_032552 | DAB2IP |
| chr2-31844 | chr2 | 1.92E+08 | 1.92E+08 | + | intron (NM_014905, intron 17 of 17) | 53761 | NM_139266 | STAT1 |
| chr2-16942 | chr2 | 1.13E+08 | 1.13E+08 | + | Intergenic | -22581 | NM_032824 | TMEM87B |
| chr4_gl000193_random-69 | chr4_gl000193_random | 78229 | 80193 | + | intron (NR_038377, intron 1 of 5) | 9164 | NR_038377 | LINC01667 |
| chr11-4802 | chr11 | 57774337 | 57774962 | + | Intergenic | -16704 | NM_001005212 | OR9Q1 |
| chr5-24563 | chr5 | 67818274 | 67819342 | + | Intergenic | 230412 | NM_001242466 | PIK3R1 |
| chr20-3684 | chr20 | 42309655 | 42310801 | + | intron (NM_001278610, intron 2 of 12) | 14569 | NM_002466 | MYBL2 |
| chr12-3716 | chr12 | 62560278 | 62560903 | + | intron (NM_178539, intron 1 of 4) | 26030 | NM_178539 | FAM19A2 |
| chr7-18981 | chr7 | 1.21E+08 | 1.21E+08 | + | Intergenic | -151085 | NM_002851 | PTPRZ1 |
| chr7-6097 | chr7 | 23355918 | 23356543 | + | intron (NM_006547, intron 12 of 14) | 17290 | NM_138446 | MALSU1 |

| | | | | | | | | |
|--------------------|-------|----------|----------|---|---------------------------------------|---------|--------------|--------------|
| chr1-19747 | chr1 | 1.02E+08 | 1.02E+08 | + | Intergenic | 193353 | NR_110212 | OLFM3 |
| chr6-8735 | chr6 | 1.05E+08 | 1.05E+08 | + | Intergenic | 676467 | NM_001321083 | HACE1 |
| chr3-3322 | chr3 | 60596379 | 60597004 | + | intron (NM_002012, intron 4 of 9) | 6912 | NR_128708 | MIR548BB |
| chr5-19765 | chr5 | 5165459 | 5167045 | + | intron (NR_136935, intron 3 of 21) | 25809 | NM_139056 | ADAMTS16 |
| chr2-32821 | chr2 | 1.34E+08 | 1.34E+08 | + | intron (NM_207481, intron 4 of 17) | 27756 | NR_110294 | LOC101928161 |
| chr12-1629 | chr12 | 1.04E+08 | 1.04E+08 | + | intron (NM_017564, intron 5 of 68) | 39931 | NM_017564 | STAB2 |
| chr3-7028 | chr3 | 1.64E+08 | 1.64E+08 | + | Intergenic | 254751 | NR_031665 | MIR1263 |
| chr18-1344 | chr18 | 5073236 | 5074119 | + | Intergenic | 123578 | NM_001145194 | AKAIN1 |
| chr14-5581 | chr14 | 66505072 | 66505832 | + | Intergenic | -447637 | NR_024338 | LINC00238 |
| chr15-4678 | chr15 | 25841774 | 25843076 | + | Intergenic | -158250 | NM_000462 | UBE3A |
| chr15-16105 | chr15 | 21925452 | 21928450 | + | Intergenic | 13788 | NR_027053 | LOC646214 |
| chr6-15798 | chr6 | 9225440 | 9226537 | + | Intergenic | 573546 | NR_004855 | HULC |
| chr2-16937 | chr2 | 1.05E+08 | 1.05E+08 | + | Intergenic | -1294 | NR_038231 | LINC01114 |
| chr13-8955 | chr13 | 99393686 | 99394649 | + | intron (NM_005073, intron 1 of 22) | 10762 | NM_005073 | SLC15A1 |
| chr12-15701 | chr12 | 6334827 | 6336318 | + | intron (NM_001330312, intron 2 of 7) | 26090 | NM_001769 | CD9 |
| chr10-16246 | chr10 | 87307577 | 87308366 | + | Intergenic | -29517 | NR_038986 | GRID1-AS1 |
| chr18-6799 | chr18 | 14747653 | 14748687 | + | promoter-TSS (NM_001145029) | -69 | NM_001145029 | ANKRD30B |
| chr5-24073 | chr5 | 49660461 | 49662186 | + | Intergenic | 75911 | NM_198449 | EMB |
| chr13-6798 | chr13 | 34083545 | 34084820 | + | intron (NM_001243476, intron 4 of 17) | 166790 | NM_001243476 | STARD13 |
| chr6-15396 | chr6 | 1.18E+08 | 1.18E+08 | + | Intergenic | 101151 | NM_138459 | NUS1 |
| chr7-22363 | chr7 | 94060023 | 94063293 | + | Intergenic | 37785 | NM_000089 | COL1A2 |

| | | | | | | | | |
|--------------------|-------|----------|----------|---|---|---------|------------------|--------------|
| chr6-11355 | chr6 | 99318686 | 99319583 | + | Intergenic | 36554 | NM_005604 | POU3F2 |
| chr4-16526 | chr4 | 14405775 | 14406827 | + | Intergenic | 292709 | NR_033931 | LINC01085 |
| chr2-33572 | chr2 | 2.25E+08 | 2.25E+08 | + | intron (NM_006216, intron 3 of 8) | 33786 | NR_073116 | SERPINE2 |
| chr2-14712 | chr2 | 1.31E+08 | 1.31E+08 | + | Intergenic | -30939 | NM_207312 | TUBA3E |
| chr1-29880 | chr1 | 91586020 | 91587158 | + | Intergenic | -98777 | NM_201269 | ZNF644 |
| chr21-4889 | chr21 | 40823179 | 40824922 | + | exon (NM_001317740, exon 1 of 6) | 270 | NM_007341 | SH3BGR |
| chr9-19547 | chr9 | 97883112 | 97884265 | + | intron (NM_001243743, intron 10 of 14) | -35312 | NR_036109 | MIR3074 |
| chr3-18661 | chr3 | 75551788 | 75553013 | + | Intergenic | -68134 | NR_024241 | FAM86DP |
| chr5-14650 | chr5 | 39506122 | 39506905 | + | Intergenic | -14020 | NR_104632 | LINC02104 |
| chr18-6639 | chr18 | 52279231 | 52280006 | + | Intergenic | 21228 | NM_173629 | DYNAP |
| chr3-10178 | chr3 | 1.63E+08 | 1.63E+08 | + | Intergenic | -12739 | NR_033945 | LINC01192 |
| chr2-12371 | chr2 | 18492473 | 18493521 | + | Intergenic | 248962 | NM_020905 | RDH14 |
| chr10-17441 | chr10 | 79595705 | 79596866 | + | intron (NM_004747, intron 7 of 31) | 90063 | NM_004747 | DLG5 |
| chr5-23205 | chr5 | 49849235 | 49850026 | + | Intergenic | -112103 | NM_0011780 55 | PARP8 |
| chr8-10137 | chr8 | 586712 | 588264 | + | intron (NM_001303100, intron 5 of 5) | -91707 | NM_175075 | TDRP |
| chr5-17189 | chr5 | 5145703 | 5146575 | + | intron (NR_136935, intron 2 of 21) | 5696 | NM_139056 | ADAMTS16 |
| chr2-10055 | chr2 | 10654682 | 10655567 | + | Intergenic | 65270 | NR_110597 | LOC101929715 |
| chr2-28398 | chr2 | 60612398 | 60614236 | + | intron (NR_132992, intron 1 of 1) | 1263 | NR_039631 | MIR4432 |
| chr3-33158 | chr3 | 1.67E+08 | 1.67E+08 | + | promoter-TSS (NR_135545) | -703 | NR_135545 | LOC105374194 |
| chr10-11519 | chr10 | 1.22E+08 | 1.22E+08 | + | Intergenic | 92061 | NR_103717 | LINC01561 |
| chr15-13330 | chr15 | 33027016 | 33033698 | + | Intergenic | -19291 | NR_109767 | LOC100131315 |
| chr12-3712 | chr12 | 53253803 | 53254555 | + | Intergenic | -11401 | NM_173352 | KRT78 |

| | | | | | | | | |
|--------------------|-------|----------|----------|---|---|--------|-----------|--------------|
| chr2-37346 | chr2 | 2.16E+08 | 2.16E+08 | + | intron (NM_001306132, intron 36 of 44) | 62002 | NM_004044 | ATIC |
| chr10-18418 | chr10 | 1217729 | 1218987 | + | Intergenic | 12650 | NR_015376 | LINC00200 |
| chr7-19799 | chr7 | 64309444 | 64310527 | + | intron (NR_073389, intron 2 of 2) | -53635 | NM_021148 | ZNF273 |
| chr3-26532 | chr3 | 83995504 | 83996259 | + | Intergenic | 922845 | NR_033860 | LINC00971 |
| chr9-5718 | chr9 | 74427132 | 74428022 | + | Intergenic | -43777 | NM_013390 | TMEM2 |
| chr1-44141 | chr1 | 2.13E+08 | 2.13E+08 | + | Intergenic | -16299 | NR_125985 | LINC01740 |
| chr2-10406 | chr2 | 1.51E+08 | 1.51E+08 | + | intron (NM_005168, intron 3 of 5) | 4391 | NM_005168 | RND3 |
| chr18-1582 | chr18 | 14978955 | 14979455 | + | Intergenic | -8738 | NR_110783 | LINC01444 |
| chr8-13786 | chr8 | 1.37E+08 | 1.37E+08 | + | intron (NM_006558, intron 6 of 8) | 142561 | NM_006558 | KHDRBS3 |
| chr20-12945 | chr20 | 11735349 | 11736553 | + | Intergenic | 115412 | NR_110635 | LINC00687 |
| chr2-6323 | chr2 | 1.23E+08 | 1.23E+08 | + | Intergenic | 147247 | NM_004622 | TSN |
| chr18-8279 | chr18 | 40511008 | 40512300 | + | intron (NM_001272077, intron 3 of 5) | 184003 | NM_002930 | RIT2 |
| chr4-9774 | chr4 | 53023877 | 53024783 | + | Intergenic | 106833 | NM_145263 | SPATA18 |
| chr3-4481 | chr3 | 1.65E+08 | 1.65E+08 | + | intron (NR_125764, intron 1 of 5) | 48096 | NR_125764 | LINC01322 |
| chr10-8723 | chr10 | 38439680 | 38440435 | + | TTS (NM_001324256) | -24542 | NR_024524 | LOC100129055 |
| chr18-4857 | chr18 | 46574400 | 46575456 | + | intron (NM_017653, intron 16 of 16) | 1210 | NR_039898 | MIR4744 |
| chr7-11570 | chr7 | 96603123 | 96604048 | + | intron (NR_015448, intron 2 of 2) | -31705 | NM_005222 | DLX6 |
| chr2-6449 | chr2 | 1.23E+08 | 1.23E+08 | + | Intergenic | 601403 | NM_004622 | TSN |
| chr18-3016 | chr18 | 40439909 | 40440544 | + | intron (NM_002930, intron 4 of 4) | 255431 | NM_002930 | RIT2 |
| chr10-22819 | chr10 | 26726611 | 26728583 | + | 5' UTR (NM_019043, exon 2 of 15) | 331 | NM_019043 | APBB1IP |
| chr9-9085 | chr9 | 71011716 | 71012995 | + | intron (NM_021965, intron 6 of 10) | -39573 | NR_121192 | PGM5-AS1 |

| | | | | | | | | |
|--------------------------|----------------|----------|----------|---|---------------------------------------|---------|--------------|--------------|
| chr3-12160 | chr3 | 84088069 | 84088825 | + | Intergenic | 830279 | NR_033860 | LINC00971 |
| chrUn_gl000218-32 | chrUn_gl000218 | 127873 | 129288 | + | Intergenic | -31126 | NR_037871 | LOC100233156 |
| chr2-37033 | chr2 | 87575441 | 87577040 | + | Intergenic | 154331 | NR_039929 | MIR4771-2 |
| chr11-27207 | chr11 | 1.01E+08 | 1.01E+08 | + | Intergenic | 173379 | NR_037485 | MIR3920 |
| chr10-14883 | chr10 | 771347 | 772367 | + | Intergenic | -36249 | NM_014974 | DIP2C |
| chr2-8432 | chr2 | 1.75E+08 | 1.75E+08 | + | intron (NR_038897, intron 1 of 2) | 1394 | NR_038897 | LINC01305 |
| chrX-9256 | chrX | 1.45E+08 | 1.45E+08 | + | Intergenic | -207588 | NM_001144004 | SLITRK2 |
| chr15-12465 | chr15 | 20398043 | 20398829 | + | Intergenic | -89561 | NR_038836 | CHEK2P2 |
| chr7-21944 | chr7 | 95592610 | 95593425 | + | intron (NM_001278422, intron 5 of 15) | 191199 | NM_001135557 | DYNC1I1 |
| chr9-20931 | chr9 | 2610700 | 2611477 | + | intron (NR_015375, intron 1 of 3) | -10686 | NM_001322226 | VLDLR |
| chr8-13418 | chr8 | 896545 | 897340 | + | intron (NM_001346810, intron 2 of 14) | 208394 | NR_134292 | LOC401442 |
| chr7-16656 | chr7 | 1.04E+08 | 1.04E+08 | + | intron (NM_199000, intron 1 of 2) | 120637 | NM_199000 | LHFPL3 |
| chrX-12157 | chrX | 1.4E+08 | 1.4E+08 | + | Intergenic | -23161 | NR_031725 | MIR320D2 |
| chr4-10265 | chr4 | 44018406 | 44019460 | + | promoter-TSS (NR_131958) | 55 | NR_131959 | LVCAT1 |
| chr9-3129 | chr9 | 18368162 | 18368787 | + | Intergenic | -105605 | NM_001040272 | ADAMTSL1 |
| chr6-22914 | chr6 | 1.65E+08 | 1.65E+08 | + | Intergenic | -212343 | NR_131926 | MEAT6 |
| chr16-6708 | chr16 | 61466302 | 61466956 | + | Intergenic | -376959 | NR_039624 | MIR4426 |
| chrX-10558 | chrX | 1.11E+08 | 1.11E+08 | + | intron (NM_014289, intron 2 of 12) | 8283 | NM_014289 | CAPN6 |
| chr10-1534 | chr10 | 25419810 | 25420435 | + | intron (NR_120649, intron 6 of 7) | 18135 | NR_120649 | LINC01516 |

| | | | | | | | | |
|--------------------|-------|----------|----------|---|--|---------|------------------|--------------|
| chr12-12611 | chr12 | 14707017 | 14707933 | + | intron (NM_024829, intron 1 of 10) | -13191 | NR_120465 | PLBD1-AS1 |
| chr5-17749 | chr5 | 1.07E+08 | 1.07E+08 | + | Intergenic | -266748 | NR_104671 | LINC01950 |
| chr8-16626 | chr8 | 88386667 | 88387548 | + | intron (NM_173538, intron 10 of 10) | 499189 | NM_152418 | DCAF4L2 |
| chr5-3445 | chr5 | 89716906 | 89717531 | + | Intergenic | 11401 | NR_105024 | LOC731157 |
| chr5-8159 | chr5 | 1.01E+08 | 1.01E+08 | + | Intergenic | -632063 | NM_005668 | ST8SIA4 |
| chr10-5294 | chr10 | 2551168 | 2552187 | + | Intergenic | -8023 | NR_136146 | LOC105376351 |
| chr21-2887 | chr21 | 28204940 | 28209546 | + | Intergenic | 10485 | NM_006988 | ADAMTS1 |
| chr20-7779 | chr20 | 7755466 | 7756285 | + | Intergenic | 165218 | NM_017545 | HAO1 |
| chr1-40426 | chr1 | 1.76E+08 | 1.76E+08 | + | Intergenic | 34562 | NR_002998 | SCARNA3 |
| chr5-5504 | chr5 | 20972503 | 20973148 | + | Intergenic | -396843 | NM_0012919 56 | CDH18 |
| chr10-14491 | chr10 | 1.34E+08 | 1.34E+08 | + | intron (NM_001318879, intron 2 of 11) | 18756 | NR_134911 | STK32C |
| chr5-21589 | chr5 | 1.7E+08 | 1.7E+08 | + | intron (NM_014592, intron 1 of 7) | 67875 | NM_0012783 39 | KCNIP1 |
| chr16-949 | chr16 | 13959918 | 13960544 | + | Intergenic | -53783 | NM_005236 | ERCC4 |
| chr18-2467 | chr18 | 14510637 | 14511262 | + | TTS (NM_001137671) | -12244 | NR_024076 | CXADRP3 |
| chr14-15264 | chr14 | 57578777 | 57579623 | + | Intergenic | -156406 | NR_026895 | AP5M1 |
| chr10-24952 | chr10 | 56318987 | 56319839 | + | intron (NM_001142771, intron 2 of 34) | 48304 | NR_031642 | MIR548F1 |
| chr9-8252 | chr9 | 2319435 | 2320395 | + | Intergenic | 161459 | NM_0012894 00 | SMARCA2 |
| chr5-2673 | chr5 | 5658311 | 5659062 | + | Intergenic | 235900 | NM_015325 | ICE1 |
| chr2-39984 | chr2 | 1.23E+08 | 1.23E+08 | + | Intergenic | 448145 | NM_004622 | TSN |
| chr14-6674 | chr14 | 21788019 | 21788912 | + | intron (NM_020366, intron 11 of 23) | 32329 | NM_020366 | RPGRIP1 |
| chr9-14787 | chr9 | 75778316 | 75792453 | + | TTS (NM_000700) | 18737 | NM_000700 | ANXA1 |
| chr1-10064 | chr1 | 17188268 | 17188768 | + | Intergenic | -3002 | NR_037446 | MIR3675 |

| | | | | | | | | |
|-------------------|------|----------|----------|---|------------------------------------|-------|------------------|--------|
| chr6-17058 | chr6 | 1.32E+08 | 1.32E+08 | + | 3' UTR (NM_001901, exon 5 of 5) | 3029 | NM_001901 | CTGF |
| chr9-20642 | chr9 | 74960852 | 74964748 | + | Intergenic | 16708 | NM_0012782 45 | ZFAND5 |

